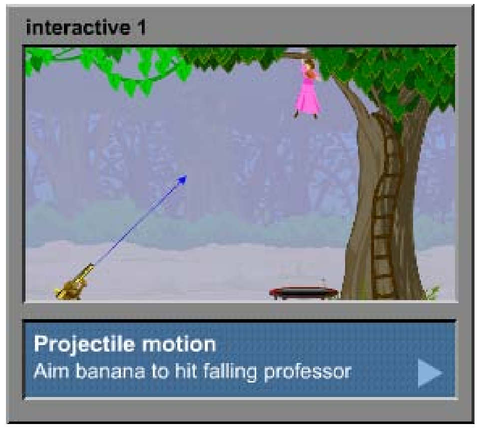
Table of Contents

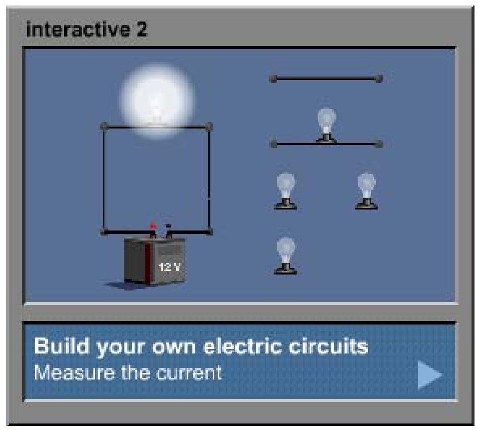
For additional information on the product(s) and purchasing, please contact Kinetic Books Company.

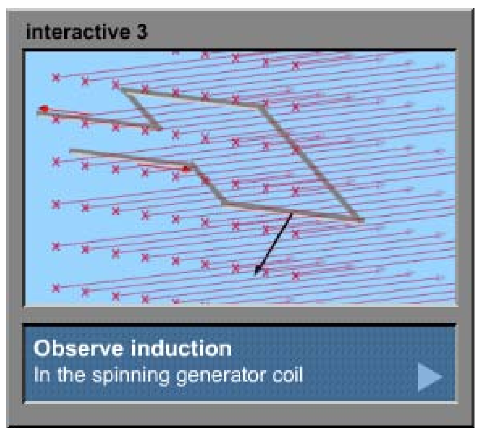
www.kineticbooks.com
1-877-4kbooks (877) 452-6657

Chapter 0 Introduction

0.0 - Welcome to an electronic physics textbook!







Textbooks, like this one, contain words and illustrations. In an ordinary textbook, the words are printed and the illustrations are static, but in this book, many of the illustrations are animations and many words are spoken. Altogether, this textbook contains more than 600,000 words, 150 simulations, 1000 animations, 5000 illustrations, 15 hours of audio narration, and 35,000 lines of Java and JavaScript code.

All this is designed so that you will experience more physics. You will race cars around curves, see the forces between charged particles, dock a space craft, generate electricity by moving a wire through a magnetic field, control waves in a string to “make music”, measure the force exerted by an electric field, and much more. These simulations and animations are designed to allow you to “see” more physics and make it easier for you to assess your learning, since many of them pose problems for you to solve.

The foundation of this textbook is the same as a traditional textbook: text like this and illustrations. Concepts like “velocity” or “Newton’s second law” are explained as they are in traditional textbooks. From there we go a step further, taking advantage of the computer and giving you additional ways to learn about physics. The textbook has many features: simulations; problems where the computer checks your answers and then works with you step by step; animations that are narrated; search capability; and much more. We will start with some simulations. In subsequent sections, we will show you how we use animations and narration to teach physics, and how a computer will help you solve problems.

At the right are three examples of how we take advantage of an interactive simulation engine. Click on any of the illustrations to start an interactive simulation; it will open a separate window. When you are done, you can close the window. The window that contains this text will remain open.

In the first simulation, you aim the monkey’s banana bazooka so that the banana will reach the professor. The instant the banana is fired, the professor lets go of the tree and falls toward the ground. You aim the banana bazooka by dragging the head of the arrow shown on the right. Aim the bazooka and then press GO. Press RESET to try again. (And do not worry: We, too, value physics professors, so the professor will emerge unscathed.)

This is an animated version of a classic physics problem and appears about halfway through a chapter of the textbook. The majority of our simulations require the calculation of precise answers, but like this one, they are all great ways to see a concept at work.

In the second simulation, you can extend a simple circuit. The initial circuit shown on the right contains a battery and a light bulb. You can add light bulbs or more wire segments by dragging them near the desired location. Once there, they will snap into place. You can also use an ammeter to measure the amount of current flowing through a section of a wire, and a voltmeter to measure the potential difference across a light bulb or the battery.

There are many experiments you can conduct with these simple tools. For instance, place a light bulb in the horizontal segment above the one which already contains the light bulb and connect it to the circuit with two additional vertical wire segments. Does this alter the power flowing through the first light bulb? The brightness of each light bulb is roughly proportional to the power the circuit supplies to it.

How do the potential differences across the light bulbs compare to one another? To the potential difference across the battery? Measure the current flowing through a piece of wire immediately adjacent to the battery, and through each of the wire segments that contains a light bulb. Do you see a mathematical relationship between these three values?

You will be asked to make observations in many simulations like this, and as you learn physics, to apply what you have learned to answer problems posed by the simulations. You will use your knowledge to do everything from juggle to dock a spaceship!

In the third simulation, you experiment with a simple electric generator. When the crank is turned, the rectangular wire loop shown in the illustration turns in a magnetic field. The straight lines you see are called magnetic field lines. Turning the handle of the generator creates an electric current and what is called an emf. The emf is measured in volts.

After you launch the simulation, you can change your point-of-view with a slider. The illustration you see to the right provides a conceptual overview of what a generator is. If you change the viewing angle, you can better see the angle between the wire and the field, and how that affects the current. A device called an oscilloscope is used to measure the emf created by the generator.

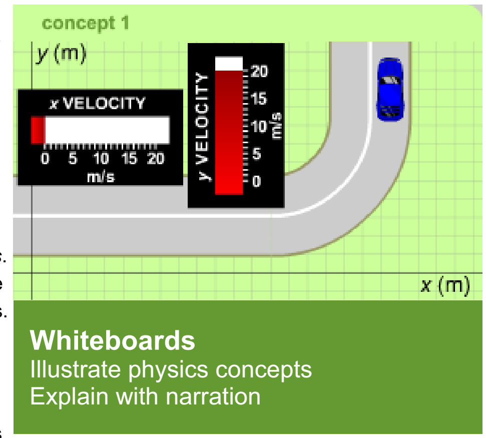
The electric generator is an advanced topic, and if you are just beginning your study of physics, it presents you with many unfamiliar concepts. However, the simulation shows how we can take advantage of software to allow you to change the viewing angle and to view processes that change over time.

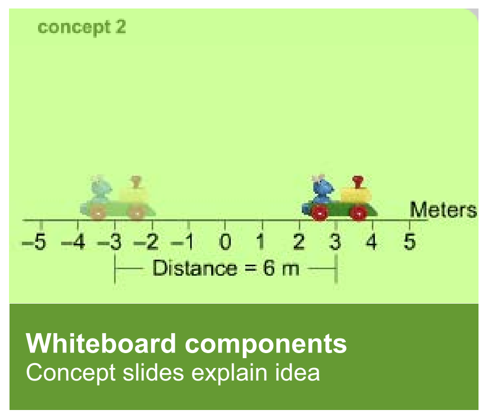
If you want to see more simulations we enjoyed creating: "dragging" a ball to match a graph, sliding a block up a plane, electromagnetic induction, electric potential, space docking mission and wave interference. You can click on any of these topics and the link will take you to that section. There are many simulations; to see even more of them, you can click on the table of contents, pick a chapter, and then click on any section whose name starts with "interactive problem."

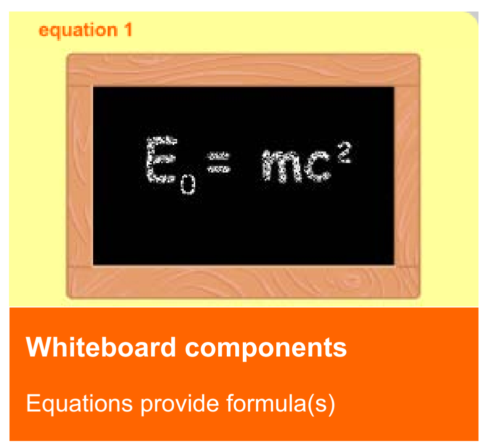
Some that address particularly sophisticated concepts include Einstein's simultaneity thought experiment, 3D views of electric flux, a mission to Mars, a wave generator used to study Fourier synthesis and phase differences in an AC circuit (complete with phasors).

To move to the next section, click on the right arrow in the black bar above or below, the arrow to the right of 0.0.

0.1 - Whiteboards







Right now, you are reading the text of this textbook. Its design is similar to that found in traditional textbooks. By "text," we mean the words you are reading and the illustrations and writing you see to the right.

As you read the words, study the illustrations and work the problems, you may feel as though you are using a traditional textbook. (We like to think it is well-conceived and well-written, but that is for you to judge.) You can print out this textbook and use it as you would a traditional print textbook.

When you use the electronic version of this book, however, you have access to an entirely different way of learning the material. It starts with what we call the *whiteboards*. You launch the whiteboards by clicking on the illustrations to the right. They present the same material discussed in the text, but do so using a sequence of narrated animations.

The text and the whiteboards cover essentially the same material. You could learn physics exclusively through the whiteboards, or you could learn it all via the words and pictures you now see. The text sometimes contains additional material: the history of the topic, an application of a principle and so on. Everything found in the whiteboards is always found in the text, so you do **not** have to click through them unless you find them a useful way to learn. The point is: You have a choice. You may also find a combination of the two particularly useful, especially for topics you find challenging.

If you are reading this on a computer, try clicking on the illustration titled "Concept 1" to the right. This will open the whiteboard in a separate window. Each whiteboard is equipped with animations, audio and its own set of controls. Both this textbook and the whiteboards can be used simultaneously. **If you do not have headphones or speakers, click on the "show text" button after you open the whiteboard**. This will allow you to read the whiteboard narration.

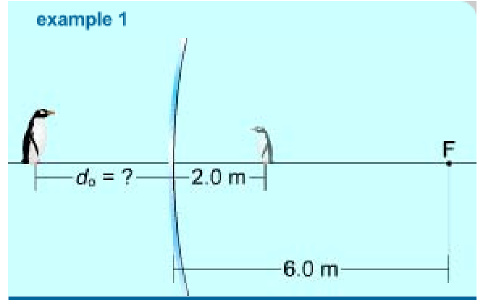
The electronic format provides a visually compelling way for you to learn what can be complex concepts and formulas. For instance, instead of a static diagram that represents a car rounding a curve, our format allows us to actually show the car moving and turning. We can also show you a greater amount of information − for example, how the horizontal and vertical velocities of the car change over time.

Typical sections throughout the book feature three graphic elements on the right side of the page, corresponding to three parts of the whiteboard. The first introduces the **concept**: For instance, what does the term "displacement" mean? The second contains the **equation**: How is displacement calculated? The third, located at the bottom right, then works an **example** problem to test your understanding of the concept and the equation.

The textbook contains hundreds of these whiteboards. If you would like to view some more to get a sense of how animation and audio play together, you can browse any chapter. You can explore topics like displacement, graphing simple harmonic motion, hitting a baseball, electric field diagrams, determining the type of image produced by a mirror and the force of a magnetic field on a moving charged particle.

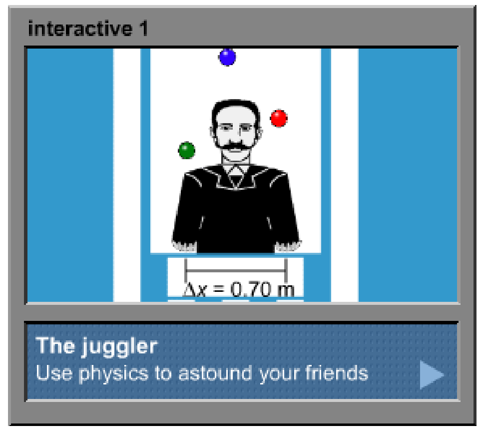
For more advanced topics: pressure-volume graphs and heat engines, spacetime diagrams and inductor-capacitor circuits.

To move to the next section, click on the right-arrow in the black bar above or below, the arrow to the right of 0.1.



Whiteboard components
Concept slides explain idea
Equations provide formula(s)
Examples work basic problems

0.2 - Interactive problems



In the first section of this chapter, we encouraged you to try various simulations. We call simulations where you set values and watch the results *interactive problems*. In this section, we explain in more detail how they work.

A sample interactive problem can be launched by clicking on the graphic on the right. When the correct x (horizontal) and y (vertical) velocities are supplied, the juggler will juggle the three balls. Before you proceed, you may wish to read the instructions below for using these interactive simulations.

As mentioned, you launch the simulations by clicking on the graphic. Typically, you will be asked to enter a value in the simulation. Sometimes you fill in a value in a text entry box, and other times you select a value using spin dials that have up and down arrow buttons.

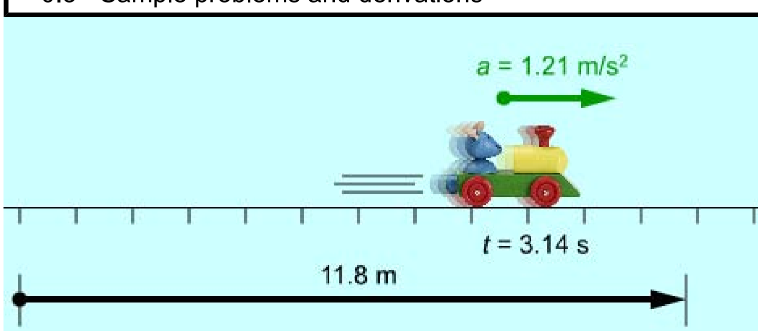
Then, you typically push the GO button and the simulation begins − things begin to move. Most simulations have a RESET button that allows you to start again. Many have a PAUSE button that makes things go three times faster (just kidding − they pause the simulation so you can record data).

Many simulations, especially those at the beginning of the chapters, just ask you to observe how entering different values changes the results. Simulations often come with gauges that display variables as they change, such as a speedometer to keep track of a car's speed as it goes around a track. You may observe the relationship between mass and the amount of gravitational force, for instance. Other simulations provide direct feedback if you succeed: the juggler juggles, you beat another racecar, and so forth.

Simulations later in the chapter often ask you to perform calculations in order to achieve a particular goal. These simulations are designed to make trial-and-error an ineffective tactic since they require a great amount of precision in the answer.

Enough preamble: Try the juggling simulation to the right. Enter any values you like for the initial y and x velocities, using the spin dials. Then press GO and watch as the juggler begins to juggle. Press RESET to enter a different set of values. A hint: One pair of values that will enable you to juggle is 6.0 m/s (meters per second) for the y velocity and 0.6 m/s for the x velocity.

0.3 - Sample problems and derivations



The mouse goes 11.8 meters in 3.14 seconds at a constant acceleration of 1.21 m/s^2.

What is its velocity at the beginning and end of the 11.8 meters?

In addition to text and interactive problem sections, this textbook contains sections with *sample problems* and *derivations* of equations. Sample problems often demonstrate a useful problem-solving technique. You see a typical sample problem above. Derivations show how an equation new to you can be created from equations you have already learned.

We follow the same sequence of steps in sample problems and derivations. (You will also follow this same sequence when you work through problems called interactive checkpoints; more on this type of problem in the next section.) Sample problems, derivations and interactive checkpoints all have some or all of the following: a diagram, a table of variables, a statement of the problem-solving strategy, the principles and equations used, and a step-by-step solution.

To show how these are organized, we work through a sample problem from the study of linear motion. The problem is stated above.

Draw a diagram

It is often helpful to draw a diagram of the problem, with important values labeled. Although almost every problem is stated using an illustration, we sometimes find it useful to draw an additional diagram.

Variables

We summarize the variables relating to the problem in a table. Some of these have values given in the problem statement or illustration. If we do not know the value of a variable, we enter the variable symbol. A variable table for the problem stated above is shown.

displacement	$\Delta x = 11.8$ m
acceleration	$a = 1.21$ m/s^2
elapsed time	$t = 3.14$ s
initial velocity	v_i
final velocity	v_f

There are two reasons we write the variables. One is so that if you see a variable with which you are unfamiliar, you can quickly see what it represents. The other is that it is another useful problem-solving technique: Write down everything you know. Sometimes you know more than you think you know! Some variables may also prompt you to think of ways to solve the problem.

After these two steps, we move to strategy.

What is the strategy?

The strategy is a summary of the sequence of steps we will follow in solving the problem. Some students who used this book early in its development called the strategy section “the hints,” which is another way to think of the strategy. There are typically many ways to solve a problem; our strategy is the one we chose to employ. (As we point out in the text when we actually solve this problem, there is another efficient manner in which to solve it.)

For the problem above, our strategy was:

1. There are **two** unknowns, the initial and final velocities, so choose **two** equations that include these two unknowns and the values you do know.
2. Substitute known values and use algebra to reduce the two equations to one equation with a single unknown value.

Principles and equations

Principles and equations from physics and mathematics are often used to solve a problem. For the problem above, for example, these two linear motion equations that apply when acceleration is constant are useful:

$$v_f = v_i + at$$

$$\Delta x = ½(v_i + v_f)t$$

The physics principles are the crucial points that the problems are attempting to reinforce. If they look quite familiar to you at some point: Great!

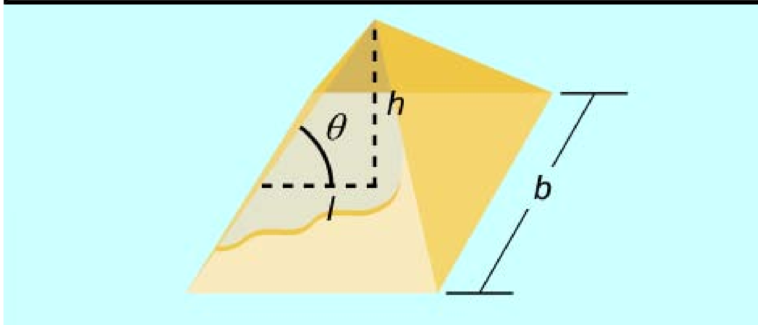
Step-by-step solution

We solve the problem (or work through the derivation) in a series of steps. We provide a reason for each step. If you want a more detailed explanation, you can click on a step, which causes a more detailed text explanation to appear on the right. Some students find the additional information quite helpful; others prefer the very brief explanation. It also varies depending on the difficulty of the problem − everyone can use a little help sometimes.

Here are the first three steps that we used to solve the problem above.

Step	Reason
1. $v_f = v_i + at$	first motion equation
2. $v_f = v_i + (1.21 \text{ m/s}^2)(3.14 \text{ s})$	substitute values
3. $v_f = v_i + 3.80 \text{ m/s}$	multiply

0.4 - Interactive checkpoints



The great pyramid of Cheops has a square base with edges that are almost exactly 230 m long. The side faces of the pyramid make an angle of 51.8° with the ground.
The apex of the pyramid is directly above the center of the base. Find its height.

This section shows you an example of an interactive checkpoint. We chose a problem that uses mathematics you may be familiar with in case you would like to solve the problem yourself.

In interactive checkpoints, all of the problem-solving elements are initially hidden. You can open any element by clicking [Show] below.

You can check your answer at any time by entering it at the top and pressing [Check] to see if you are right. You will find this is far more efficient than keying all the information into the computer, which provides a good motivation for you to solve the problem yourself. However, if you are stuck, you can always have the computer help you.

In the parts called Variables, Strategy, and Physics principles and equations, the computer will show you the information you need when you ask. In the Physics principles and equations section, we show you the principles and equations you need to solve the problem, as well as some that do not apply directly to the problem.

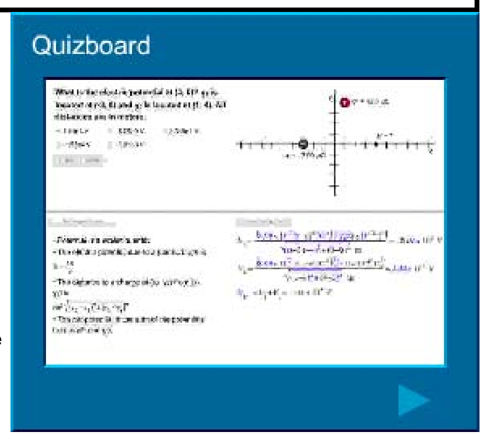
In the Step-by-step solution, you choose from the equations by clicking on the one you think you need to use. You must enter the correct values in each Step-by-step part of the solution to proceed to the next step.

Answer:

h = ______ m

0.5 - Quizboards

Each chapter has a quizboard containing several multiple-choice conceptual and quantitative problems. There are over 250 quizboard problems throughout the textbook. Quizboards allow you to test your understanding of a chapter. You see a quizboard on the right.



Quizboards appear between the summary section and the problems section in every chapter. The quizboard for a chapter can be launched from the quizboard section by clicking on the image on the right side of the page. The quizboards are designed to enable you to review many of the crucial ideas in a chapter.

Each problem in a quizboard consists of four parts: the question, the answer choices, the hints, and the solution. If you think you know the answer to the problem, choose it and click the "Check answer" button. A message will appear telling you whether you are correct. You can keep trying until you get the problem right. If you are having trouble, click "Give me a hint". Every time you click this button, a new hint appears until there are no more hints available. You can always click "Show solution" if you find yourself completely stuck.

Use the "next" and "previous" buttons on the gray bar at the bottom of the window to navigate between problems. You do not have to answer a problem correctly before moving on, so you can skip problems and come back to them later. If you use the "previous" button to go back to a problem, it will appear unanswered (even if you answered it before) so that you can try the problem again.

Click on the image to the right to use a sample quizboard. You do not need to know any physics to answer these questions. Good luck!

0.6 - Highlighting and notes

You can add notes or highlight text on most sections of the textbook. Notes always appear at the top of the section. You can use a note to write short messages about key elements of a section, or to remind yourself not to forget the extra soccer practice or to pick up the groceries.

As the note above says, you insert notes by pressing Add Note at the bottom of the page. You remove a note by clicking on the Delete button located next to the note. Modify a note by pressing the Edit button.

The text you are reading now is highlighted. To highlight text, click the Highlight button at the bottom of the page to switch it from "Off" to "On". When it is "On", any text you select (by clicking on your mouse and dragging) will be highlighted. You can remove all highlighting from a section by pressing the "Clear" button.

If the text above is **not** highlighted, your operating system or browser does not enable us to offer this feature. For instance, the feature is not available on the Macintosh operating system 0S X 10.2.

If you use a shared computer, the highlighting and notes features may be turned off. The preferences page allows you to enable or disable either of these features. You will find the preferences page by clicking on the Preferences button at the bottom of the page.

Notes and highlighting are not supported on our Web Access option or the trial version of the product on our web site.

0.7 - Online Homework

This textbook was designed to support online assessment of homework. Instructors can assign specific problems online, and you submit your responses over the Internet to a central computer. You can work offline and submit the answers when you are ready.

The computer checks the answers, and sends a report about your efforts, and the efforts of your peers, to your instructor. Your instructor can configure this service in a variety of fashions. For instance, they can set deadlines for homework assignments or decide if you are allowed to try answering a question a few times.

Online Homework is an optional feature; not all instructors will use it. If your instructor has supplied you with a login ID or told you to sign up for Online Homework, please log in now. If you are unsure, please check with your instructor.

If you want to learn more about on-line assessment in general, click here.

0.8 - Finding what you need in this book

You can navigate through the book using the Table of Contents button. When you roll your mouse over it, you will see three links.

The Chapter TOC link takes you to the table of contents for the chapter you are currently in. You will see more sections than you might see in a typical physics textbook table of contents. We chose to make it very easy to navigate to each element of the textbook by listing sample problems, derivations and other elements discretely.

The Main TOC link takes you to the list of all the chapters in the textbook. Clicking on the third link, Physics Factbook, opens a reference tool containing useful information including mathematics review topics and formulas, unit conversion factors, fundamental physical constants, properties of the elements, astronomical data, and physics equations. The Factbook also has a built-in search feature to help you find information quickly.

This textbook has no index, but likely you will find that entering text in the “search box” is more useful. Search is located at the bottom of each Web page. Search performs its task by looking at the name of each section, at the first (or essential) time any term is defined, and at some other types of text. Typing in a phrase like “kinetic energy” will produce a number of useful results.

When you use search, you do not have to worry about sequence: You do not have to guess whether we listed something under, say, “average velocity” or “velocity average.” Search looks for the terms and presents them to you along with some of their context.

That is it for logistics. The people who worked on this textbook − about 50 of us − hope you enjoy it. We have a passion for physics, and we hope some of that carries on to you.

To explore the rest of the book, move your mouse over a Table of Contents button at the top or bottom of this page, and select the Main TOC.

Chapter 1 Measurement and Mathematics

1.0 - Introduction

Heavyweight, lightweight, overweight, slender. Small, tall, vertically impaired, “how’s the weather up there?” Gifted, average, 700 math/600 verbal, rocket scientist. Gorgeous, handsome, hunk, babe.

Humans like to measure things. Whether it is our body size, height, IQ or looks, everything seems to be fair game.

Physics will teach you to measure even more things. For example, quantities such as displacement, velocity and acceleration are crucial to understanding motion. Other topics have yet more things to quantify: Mass and period are concepts required to understand the movement of planets; resistance and current are used for analyzing electric circuits. Just as you have developed a vocabulary for the things you measure, so have physicists.

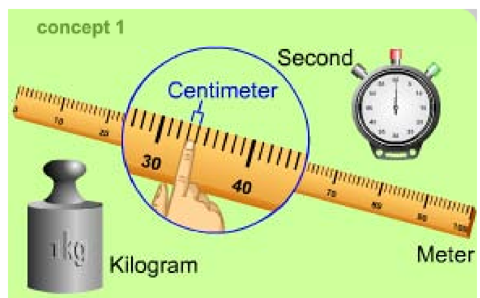
There are many different units for measuring different properties. It is possible to go all the way from A through Z in units: amperes, bars, centimeters, dynes, ergs, farads, grams, hertz, inches, joules, kilograms, liters, meters, newtons, ohms, pascals, quintals, rydbergs, slugs, teslas, unit magnetic poles, volts, webers, x units, years, and zettabars. (OK, we had to stretch for X, but it is a real unit.)

Physicists have so many units of measure at their disposal because they have plenty to measure. Physicists use amperes to tell how much electric current flows through a wire, “pascals” quantify pressure, and “teslas” are used to measure the strength of a magnetic field. If you so desired, you could become a units expert and impress (or worry) your classmates by casually noting that the U.S. tablespoon equals 1.04 Canadian tablespoons, or deftly differentiating between the barrel, U.K. Wine, versus the barrel, U.S. federal spirits, or the barrel, U.S. federal, all of which define slightly different volumes. Or you could become an international sophisticate, telling friends that one German *doppelzentner* equals about 77,162 U.K. *scruples,* which of course equals approximately 101.97 metric *glugs,* which comes out to 3120 *ukies,* a Libyan unit used for the sole purpose of measuring ostrich feathers and wool.

Fortunately, you do not need to learn units such as the ones mentioned immediately above, and you will learn the others over time. Textbooks like this one provide tables that specify the relationships between commonly used units and you will use these tables to convert between units.

1.1 - The metric system and the Système International d’Unités

Metric system: The dominant system of measurement in science and the world.



Metric system and the *Système International*
System defines fundamental units
Larger/smaller units based on powers of 10

Historically, people chose units of measure related to everyday life (the “foot” is one example). Scientists continued this tradition, developing units such as “horsepower” to measure power.

The French challenged this philosophy of measurement during their Revolution, when they decided to give measurement a more scientific foundation. Instead of basing their system on things that change − the length of a person’s foot changes during her lifetime, for example − the French based their system on what they viewed as constant. To accomplish this, they created units such as the meter, which they defined as a certain fraction of the Earth’s circumference. (To be specific: one ten-millionth of the meridian passing through Paris from the equator to the North Pole. It turns out that the distance from the equator to the North Pole does vary, but the metric system’s intent of consistency and measurability was exactly on target.)

The metric system is also based on another inspired idea: units of measurement should be based on powers of 10. This differs from the British system, which provides more variety: 12 inches in a foot, 5280 feet to a mile and so forth.

The metric system makes conversions much simpler to perform. For example, in order to calculate the number of inches in a mile, you would typically multiply by 5280 (for feet in a mile) and then by 12 (for inches in a foot). However, in the metric system, to convert between units, you typically multiply by a power of 10. For instance, to convert from kilometers to meters, you multiply by 1000. The prefix “kilo” means 1000.

The revolutionaries were a little extreme (as revolutionaries tend to be) and they held onto their position of power for only a decade or so. While some of their legacy (including their political art, rather mediocre as is much political art) has been forgotten, their clever and sensible metric system endures. Most scientists, and most countries, use the metric system today.

Scientists continue to update and refine the metric system. This expanded and updated system of measurement used today is called the *Système International d’Unités*, or SI. We typically use SI units in this textbook; several times, though, we refer to different units that may be better known to you or are commonly used in the sciences. We will discuss some of the SI units further in this chapter.

Over the years, scientists have refined measurement systems, making the definition of units ever more precise. For example, instead of being based on the Earth's circumference, the meter is now defined as the distance light travels in a vacuum during the time interval of 1/299,792,458 of a second. Although perhaps not as memorable as the initial standard, this definition is important because it is constant, precise, indestructible, and can be reproduced in laboratories around the world.

In addition to using meters for length, the *Système International* uses seconds (time), kilograms (mass), amperes (electric current), kelvins (temperature), moles (amount of substance) and candelas (luminous intensity). Many other *derived units* are based on these fundamental units. For instance, a newton measures force and is equal to kilograms times meters per second squared. On Earth, the force of gravity on a small apple is about one newton.

At the risk of drowning you in terminology, we should point out that you might also encounter references to the MKS (meter/kilogram/second) and CGS (centimeter/gram/second) systems. These systems are named for the units they use for length, mass and time.

1.2 - Prefixes

Metric units often have prefixes. Kilometers and centimeters both have prefixes before the word "meter." The prefixes instruct you to multiply or divide by a power of 10: **kilo** means multiply by 1000, so a kilometer equals 1000 meters. **Centi** means divide by 100, so a centimeter is one one-hundredth of a meter. In other words, there are 100 centimeters in a meter. The table in Equation 1 on the right lists the values for the most common prefixes.

Prefixes allow you to describe the unimaginably vast and small and everything in between. To illustrate, every day the City of New York produces 10 *giga*grams of garbage. The distance between transistors in a microprocessor is less than a *micro*meter. The power of the Sun is 400 *yotta*watts (a yotta corresponds to the factor of 10^{24}). It takes 3.34 *nano*seconds for light to travel one meter. The electric potential difference across a nerve cell is about 70 *milli*volts.

These prefixes can apply to any unit. You can use gigameters to conveniently quantify a vast distance, gigagrams to measure the mass of a huge object, or gigavolts to describe a large electrical potential difference.

Some of the most common prefixes − kilo, mega, and giga − are commonly used to describe the specifications of computers. The speed of a computer microprocessor is measured by how many computational cycles per second it can perform. Microprocessor speeds used to be specified in *mega*hertz (one million cycles per second) but are now specified in *giga*hertz (one billion cycles per second). Modem speeds have increased from *kilo*bits to *mega*bits per second. (Although bits are not part of the metric system, computer scientists use the same prefixes.)

The units of measurement you use are a matter of both convenience and convention. For example, snow skis are typically measured in centimeters; a ski labeled "170" is 170 centimeters long. However, it could also be called a 1.7-meter ski or a 1700-millimeter ski. The ski industry has decided that centimeters are reasonable units and has settled on their use as a convention.

In this textbook, you are most likely to encounter kilo, mega and giga on the large side of things and centi, milli, micro and nano on the small. Some other prefixes are not as common because they just do not seem that useful. Is it easier to say "a decameter" than the more straightforward 10 meters? And, for the extremely large and small, scientists often use another technique called scientific notation rather than prefixes.

concept 1

Prefixes
Create larger, smaller units

equation 1

prefix	symbol	factor	prefix	symbol	factor
tera	T	10^{12}	deci	d	10^{-1}
giga	G	10^{9}	centi	c	10^{-2}
mega	M	10^{6}	milli	m	10^{-3}
kilo	k	10^{3}	micro	μ	10^{-6}
hecto	h	10^{2}	nano	n	10^{-9}
deka	da	10^{1}	pico	p	10^{-12}

Prefixes
Common prefixes for powers of 10

example 1

What is the distance between the towns in kilometers?

1000 m = 1 km

$$(5000\,\cancel{\text{m}})\left(\frac{1\text{ km}}{1000\,\cancel{\text{m}}}\right) = 5\text{ km}$$

1.3 - Scientific notation

Scientific notation: A system, based on powers of 10, most useful for expressing very large and very small numbers.

Physicists like to measure the very big, the very small and everything in between. To express the results efficiently and clearly, they use scientific notation.

Scientific notation expresses a quantity as a number times a power of 10. Why is this useful? Here's an example: the Earth is about 149,000,000,000 meters from the Sun. You could express that distance as we just did, with a long string of zeros, or you could use scientific notation to write it as 1.49×10^{11} meters. The latter method has proven itself to be clearer and less prone to error.

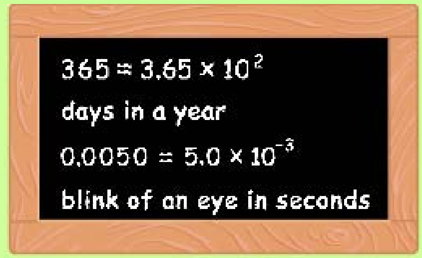
The value on the left (1.49) is called the *leading value*. The power of 10 is typically chosen so the leading value is between one and 10. In the example immediately above, we multiplied by 10^{11} so that we could use 1.49. We also could have written 14.9×10^{10} or 0.149×10^{12} since all three values are equal, but a useful convention is to use a number between one and 10.

In case you have forgotten how to use exponents, here's a quick review. Ten is the *base number*. Ten to the first power is 10; 10^2 is ten to the second (ten squared) or 100; ten to the third is 10 times 10 times 10, or 1000. A positive exponent tells you how many zeros to add after the one. When the exponent is zero, the value is one: 10^0 equals one.

As mentioned, scientists also measure the very small. For example, a particle known as a muon has a mean lifespan of about 2.2 millionths of a second. Scientific notation provides a graceful way to express this number: 2.2×10^{-6} (2.2 times 10 to the negative sixth). To review the mathematics: ten to the minus one is 1/10; ten to the minus two is 1/100; ten to the minus three is 1/1000, and so forth.

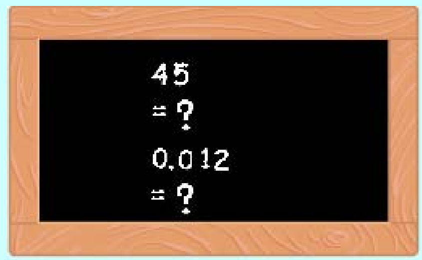
You can also write 1.49×10^{11} as 1.49e11. The two are equivalent. You may have seen this notation in computer spreadsheet programs such as *Microsoft® Excel*. We do not use this "e" notation in the text of the book, but if you submit answers to homework problems or interactive checkpoints, you will use it there.

concept 1



Scientific notation
Number between 1 and 10 (leading value)
Multiplied by power of 10

example 1



How do you write the numbers above in scientific notation?

$45 = 4.5 * 10 = 4.5*10^1$

$0.012 = 1.2 * 1/100 = 1.2*10^{-2}$

1.4 - Standards and constants

Standard: A framework for establishing measurement units.

Physical constant: An empirically based value.

Physicists establish standards so they can measure things consistently; how they define a standard can change over time. For example, the length of a meter is now based on how far light travels in a precise interval of time. This replaces a standard based on the wavelength of light emitted by krypton-86. Prior to that, the meter was defined as the distance between enscribed marks on platinum-iridium bars. Advances in technology, and the requirement for increased precision, cause scientists to change the method used to define the standard. Scientists strive for precise standards that can be reproduced as needed and which will not change.

By choosing standards, scientists can achieve consistent results around the globe and compare the results of their experiments. Well-equipped labs can measure time using atomic clocks like the one shown in Concept 1 on the right. These clocks are based on a characteristic frequency of cesium atoms. You can access the official time, as maintained by an atomic clock, by clicking here.

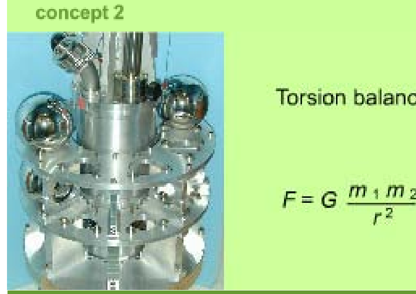
You will encounter two types of constants in this textbook. First, there are mathematical constants like π or the number 2. Second, there are physical constants, such as the gravitational constant, which is represented with a capital G in equations. We show its value in Concept 2 on the right. Devices such as the torsion balance shown are used to gather data to determine the value of G. This is an active area of research, as G is the least precisely known of the major physical constants.

Constants such as G are used in many equations. G is used in Sir Isaac Newton's law of gravitation, an equation that relates the attractive force between two bodies to their masses and the square of the distance between them. You see this equation on the

concept 1

Standards
Establish benchmarks for measurement

concept 2



Physical constants
Empirically determined values

right.

G = 6.674 2×10^{-11} N·m^2/kg^2

1.5 - Length

If you live in a country that uses the metric system, you already have an intuitive sense of how long a meter is. You are likely taller than one meter, and probably shorter than two. If you are a basketball fan, you know that male professional basketball centers tend to be taller than two meters while female professional centers average about two meters.

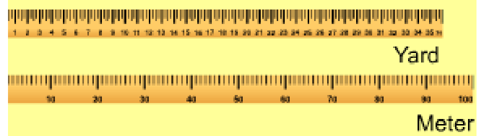
If you live in a country, such as the United States, that still uses the British system, you may not be as familiar with the meter. A meter equals about 3.28 feet, or 39.4 inches, which is to say a meter is slightly longer than a yard. The kilometer is another unit of length commonly used in metric countries. You may have noticed that cars often have speedometers that show both miles per hour and kilometers per hour. A kilometer (1000 meters) equals about 0.621 miles. In track events, a metric mile is 1.50 kilometers, which is about 93% of a British mile.

Centimeters (one one-hundredth of a meter) are also frequently used metric units. One inch equals 2.54 centimeters, so a centimeter is about four tenths of an inch. One foot equals 30.48 centimeters. You see some common abbreviations in Equation 1 to the right: “m” for meters, “km” for kilometers and “cm” for centimeters.

concept 1

Length
Measured in meters (m)
Distance light travels in
3.335 640 95×10^{-9} seconds

equation 1



Length
1 meter (m) = 3.28 feet
1 kilometer (km) = 0.621 miles
1 centimeter (cm) = 0.394 inches

1.6 - Time

Despite the appeal of measuring a lifetime in daylights, sunsets, midnights, cups of coffee, inches, miles, laughter, or in strife, physicists still choose “seconds.” How refreshingly simple!

However, as you might expect, physicists have developed a precise way to define a second. Atomic clocks, such as the one shown in Concept 1 to the right, rely on the fact that cesium-133 atoms undergo a transition when exposed to microwave radiation at a frequency of 9,192,631,770 cycles per second. These clocks are extremely accurate. Thousands of years would pass before two such clocks would differ even by a second. If you are an exceedingly precise person, you might want to consider buying a wristwatch that calibrates itself via radio signals from an atomic clock. For now, though, you can visit a web site that displays the current time as measured by an atomic clock.

In addition to being used to measure a second, atomic clocks are used to keep time. The length of a day on Earth, measured by the time to complete one rotation, is not constant. Why? The frictional force of tides causes the Earth to spin more slowly. This means that the day is getting longer (does it not just feel that way sometimes?). Every fifteen months or so since 1978, a leap second has been added to official time-keeping clocks worldwide to compensate for increased time it takes the Earth to complete a revolution.

concept 1

Time
Measured in seconds (s)
1 second = 9,192,631,770 cycles

equation 1

Time
1 hour = 3600 seconds
1 day = 86,400 seconds

1.7 - Mass

concept 1

Mass
Measured in kilograms (kg)
Resistance to change in motion
Not weight!

The standard unit of mass is the *kilogram*. (The British system equivalent is the *slug*, which is perhaps another reason to go metric.)

Physicists define mass as the property of an object that measures its resistance to a change in motion. A car has more mass than a bicycle. The three people shown straining at the car above will cause it to accelerate slowly; if they were pushing a bicycle instead, they could increase its speed much more quickly. Once they do set the car in motion, if they are not careful, its mass might prevent them from stopping it.

The official kilogram, the International Prototype Kilogram, is a cylinder of platinum-iridium alloy that resides at France's International Bureau of Weights and Measures. Copies of this kilogram reside in other secure facilities in different countries and are occasionally brought back for comparison to the original.

A liter of water has a mass of about one kilogram. A typical can of soda contains about 354 milliliters and has a mass of 0.354 kilograms.

It is tempting to write that one kilogram equals about 2.2 pounds, but this is wrong. The pound is a unit of weight; kilograms and slugs are units for mass. Weight measures the force of gravity that a planet exerts on an object, while mass reflects that object's resistance to change in motion. A classic example illustrates the difference: Your mass is the same on the Earth and the Moon, but you weigh less on the Moon because it exerts less gravitational force on you. On Earth, the force of gravity on one kilogram is 2.2 pounds but the force of gravity on a kilogram is only 0.36 pounds on the Moon.

Kilogram is abbreviated as kg. We typically use kilograms in this book, not grams (which are abbreviated as g).

The three units you need in order to understand motion, force and energy, the topics that start a physics textbook, are meters, kilograms and seconds. Other units used in studying these topics are derived from these fundamental units.

equation 1

Mass
One kilogram = one liter of water
One gram = about 25 raindrops

1.8 - Converting units

At times, you will need to convert units. Some conversion factors you know, such as 60 seconds in a minute, 12 inches in a foot, etc. Others, such as the number of seconds in a year, require a bit of calculation.

Keep in mind that if you do not use consistent units, troubles will arise. NASA dramatically illustrated the cost of such errors when it lost a spacecraft in 1999. A company supplied data to NASA based on British units (pounds) when NASA engineers expected metric units (newtons). Oops. That, alas, was the end of that space probe (and about $125 million and, one suspects, some engineer's NASA career).

Speedometers often show speeds in both mi/h and km/h

As NASA's misfortune indicates, you need to make sure you use the correct units when solving physics problems. If a problem presents information about a quantity like time in different units, you need to convert that information to the same units.

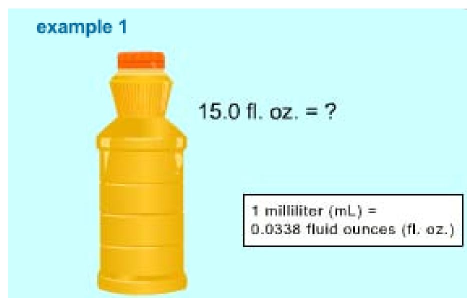
You convert units by:

1. Knowing the conversion factor (say, 12 inches to a foot; 2.54 centimeters to an inch; $125 million to a spacecraft).
2. Multiplying by the conversion factor (such as 3.28 feet/1.00 meter) so that you cancel units in both the numerator and denominator. For example, to convert meters to feet, you multiply by 3.28 ft/m so that the meter units cancel. This may be easier to understand by viewing the example on the right.

In conversions, it is easy to make mistakes so it is good to check your work. To make sure you are applying conversions correctly, make sure the appropriate units cancel. To do this, you note the units associated with each value and each conversion factor.

As is shown on the right, a unit that is in both a denominator and a numerator cancels. You should look to see that the units that remain "uncancelled" are the ones that you desired. For instance, in the example problem, fluid ounces cancel out, and the desired units, milliliters, remain.

Converting units
Choose appropriate conversion factor
Multiply by conversion factor as a fraction
Make sure units cancel!



How many milliliters of orange juice in the bottle?

1 mL = 0.0338 fl. oz.

Conversion fraction: $\dfrac{1\ \text{mL}}{0.0338\ \text{fl. oz.}}$

$$15.0\ \cancel{\text{fl. oz.}} \times \frac{1\ \text{mL}}{0.0338\ \cancel{\text{fl. oz.}}} = 444\ \text{mL}$$

1.9 - Interactive problem: converting units

For the red car to go fast enough to win the race − but not so fast as to skid off the track − it needs to travel at 24.6 miles per hour (mi/h). The car's control panel only accepts meters per second (m/s) as input, however. You need to convert from mi/h to m/s.

Some conversion factors that may help include 5,280 feet to a mile and 0.305 meters to a foot.

In the simulation, enter the speed in m/s, to the nearest 0.1 m/s. Press GO to start the race. Press RESET if you want to try again.

If you have trouble, consult the section on converting units.

1.10 - Dimensional analysis

An example of a physical measurement is “3.0 meters.” The meter is a unit, the type of unit being determined by the *dimension* to be measured. Meters are units of length, as are kilometers, feet, miles, inches and so on. Whichever of these units is used, length is the dimension being measured. Other fundamental dimensions include time (seconds, hours, days and so on) and mass (kilograms, grams).

Dimensional analysis is a useful tool for analyzing physical situations and checking whether calculations make sense. In dimensional analysis, dimensions are treated algebraically. We use the symbols L, T, and M to represent the dimensions of length, time, and mass. The volume of a cube, for instance, has dimensions L×L×L or L^3.

An example will demonstrate the usefulness of dimensional analysis. First, we introduce several terms from the study of motion. The average velocity v of an object is its displacement (the net distance it moves) divided by the time it travels. The dimensions of velocity, then, are L/T. An object's acceleration a, on the other hand, has dimensions L/T^2.

Let’s say you roughly recall an equation for calculating the distance, x, that an object that starts at rest travels in time t, but cannot remember for sure if it uses velocity or acceleration. But you are sure the equation is either

$x = ½vt^2$ or $x = ½at^2$

You can determine the correct equation using dimensional analysis. The left side of the equation, x, represents the distance traveled and has dimension L. So the right side of the equation must also have dimension L. You can then check to see that

dimension of $½vt^2 = (L/T)(T^2) = LT$

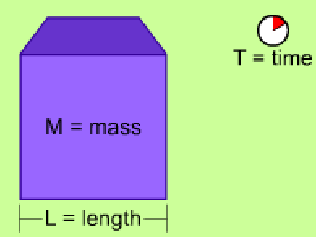
but

dimension of $½at^2 = (L/T^2)(T^2) = L$

So the correct form of the equation is the one using the acceleration a.

Notice that we ignored the ½ in the equation. Some quantities used in physics equations are *dimensionless*, meaning they have no dimension and do not carry units. All pure numbers like ½ or π are dimensionless, as are some measures of angles, like π radians.

concept 1



Dimensions
Attributes like length, time, mass
Combine algebraically
· Volume of cube has dimension L^3

equation 1

$$x = \frac{1}{2}at^2 \qquad \frac{1}{d_i} + \frac{1}{d_o} = \frac{1}{f}$$
$$F = \frac{Gm_1m_2}{r^2} \qquad L = mvr$$
$$Q = cm\Delta T \qquad a_c = \frac{v^2}{r}$$
$$P = I^2R$$

Dimensions
For valid equations:

left dimension = right dimension

example 1

$$\frac{Gm_1m_2}{r^2}$$

Dimension of G = "L^3/MT^2"

***G* is a constant with the dimension shown; m_1 and m_2 measure mass; and *r* is a**

distance. What is the dimension of the expression?

$$\text{Dimension} = \frac{(L^3/MT^2)(M)(M)}{L^2}$$

$$\text{Dimension} = \frac{ML}{T^2}$$

1.11 - Multiplying numbers in scientific notation

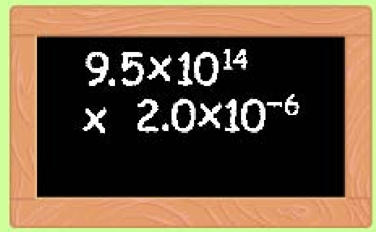
Scientific notation makes it simpler to multiply and divide large and small numbers. In this section, we focus on multiplication.

When multiplying numbers expressed using scientific notation, you first find the leading value of the result by multiplying the leading values of the numbers in the traditional way. If you want to multiply 2×10^3 by 3×10^4, the leading values are two and three, and their product is six. This is the leading value of the product.

Next, **add** the exponents to find the power of ten. In the example above, the exponents are three and four. You can add them to get the result 10^7. The final result of the multiplication is 6×10^7. The principles of how to multiply numbers in scientific notation and an example problem are shown to the right.

With scientific notation, you should format your number after the multiplication so that the leading value remains between one and 10. In other words, if you multiply 3.0 by 4.0, you should report the value as 1.2×10^1, not 12×10^0.

concept 1



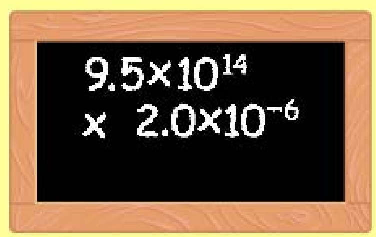
Multiplying numbers in scientific notation

Multiply the leading values

Add the exponents

· Adjust so leading value is between 1 and 10

equation 1



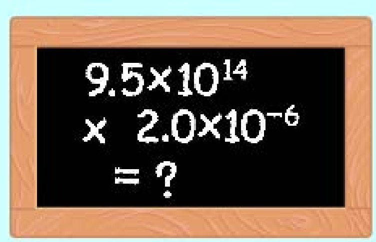
Multiplying numbers in scientific notation

$$(x\times10^{a})(y\times10^{b}) = xy\times10^{a+b}$$

x, y = leading values

a, b = exponents

example 1



What is the product of the two numbers?

$xy\times10^{a+b}$

$(9.5\times10^{14})(2.0\times10^{-6})$
$9.5\cdot2.0\times10^{14-6}$
19×10^{8}
1.9×10^{9}

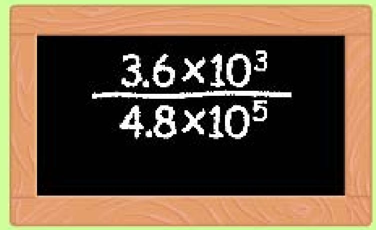
1.12 - Dividing numbers in scientific notation

In this section, we discuss how to divide numbers expressed in scientific notation. When dividing, you first find the leading value of the result by dividing the leading values in the traditional manner. For example, if you want to divide 6×10^4 by 3×10^2, you first divide six by three, with a result of two.

Then, you subtract the exponents to find the power of 10. In the example above, the exponents are four and two, so the power of 10 is 10^2. The final result of the division is 2×10^2. The principles of how to divide in scientific notation and an example problem are shown on the right.

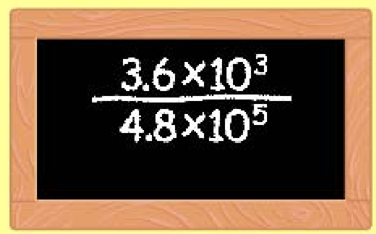
Remember to format your number after the division so that the leading value remains between one and 10. In other words, if you divide two by four, you should report the value as 5×10^{-1}, not 0.5×10^0.

concept 1



Dividing numbers in scientific notation
Divide the leading values
Subtract the exponents
· Adjust so leading value is between 1 and 10

equation 1

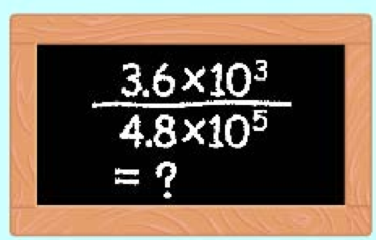


Dividing numbers in scientific notation

$(x\times10^{a})/(y\times10^{b}) = (x/y)\times10^{a-b}$

x, y = leading values
a, b = exponents

example 1



What is the result of dividing the two numbers?
$(x/y)\times10^{a-b}$
$(3.6\times10^{3})/(4.8\times10^{5})$
$(3.6/4.8)\times10^{3-5}$
0.75×10^{-2}

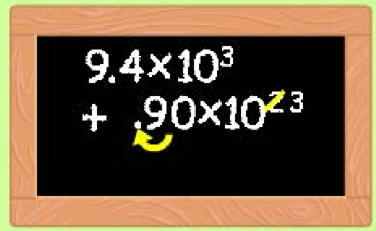
7.5×10^{-3}

1.13 - Adding and subtracting numbers in scientific notation

When two numbers in scientific notation have different powers of 10, you must perform an important step before adding and subtracting them: you must first establish a common power of 10 for both numbers. If you want to add 1.20×10^1 and 2.40×10^2, you cannot add 1.20 and 2.40 to yield 3.60. Why? Because the two numbers you are adding are 12.0 and 240, which sum to 252, not 36 or 360.

In this example, the numbers in scientific notation have different powers of 10: the first has 10 to the first power, the second, 10 squared. This means you must first convert to a common power of 10. For example, convert the first number to 0.120×10^2 and then add to obtain 2.52×10^2.

concept 1



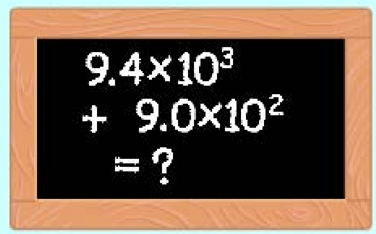
Adding and subtracting numbers in scientific notation

Adjust so exponents match

Then add or subtract leading values only

· Adjust so leading value is between 1 and 10

example 1



What is the sum of the two numbers?

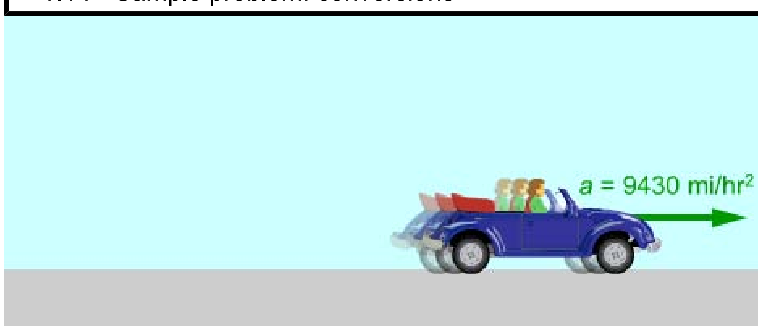
$9.4\times10^3 + 9.0\times10^2$

$9.4\times10^3 + 0.9\times10^3$

10.3×10^3

1.03×10^4

1.14 - Sample problem: conversions



Express the car's acceleration in m/s², using scientific notation. For an extra challenge, state the result in terameters per second squared.

Variables

We will use a to represent the acceleration we are converting.

What is the strategy?

1. Express the car's acceleration in scientific notation.
2. Using conversion factors in scientific notation, convert miles to yards and then to meters.
3. Convert hours squared to minutes squared and then to seconds squared.
4. Go for the challenge! Convert to terameters.

Conversion factors and prefixes

1 mi = 1760 yd

1 yd = 0.914 m

1 h = 60 min

1 min = 60 s

tera = 10^{12}

Step-by-step solution

We first write the acceleration in scientific notation and convert miles to yards to meters.

Step	Reason
1. $a = 9430\ \text{mi/h}^2 = 9.43\times10^3\ \text{mi/h}^2$	scientific notation
2. $a = \left(9.43\times10^3\ \frac{\cancel{\text{mi}}}{\text{h}^2}\right)\left(\frac{1.76\times10^3\ \text{yd}}{1\ \cancel{\text{mi}}}\right)$ $a = 1.66\times10^7\ \frac{\text{yd}}{\text{h}^2}$	multiply by factor converting miles to yards
3. $a = \left(1.66\times10^7\ \frac{\cancel{\text{yd}}}{\text{h}^2}\right)\left(\frac{9.14\times10^{-1}\ \text{m}}{1\ \cancel{\text{yd}}}\right)$ $a = 1.52\times10^7\frac{\text{m}}{\text{h}^2}$	multiply by factor converting yards to meters

Now we convert hours squared to minutes squared to seconds squared. Because the units we are converting are squared, we square the conversion factors. As requested, we state the final result in scientific notation.

Step	Reason
4. $a = \left(1.52\times10^7\ \frac{\text{m}}{\cancel{\text{h}}^2}\right)\left(\frac{1\ \cancel{\text{h}}}{6.0\times10^1\ \text{min}}\right)^2$ $a = 4.22\times10^3\ \frac{\text{m}}{\text{min}^2}$	multiply by square of factor converting hours to minutes
5. $a = \left(4.22\times10^3\ \frac{\text{m}}{\cancel{\text{min}}^2}\right)\left(\frac{1\ \cancel{\text{min}}}{6.0\times10^1\ \text{s}}\right)^2$ $a = 1.17\times10^0\ \text{m/s}^2$	multiply by square of factor converting minutes to seconds

Finally, we convert to terameters per second squared, or Tm/s^2. "Tera" means 10^{12}.

Step	Reason
6. $a = \left(1.17\times10^0\ \frac{\cancel{\text{m}}}{\text{s}^2}\right)\left(\frac{1\ \text{Tm}}{10^{12}\ \cancel{\text{m}}}\right)$ $a = 1.17\times10^{-12}\ \text{Tm/s}^2$	multiply by conversion factor for terameters

1.15 - Interactive checkpoint: conversions

Farmer Betty's farm covers 3570 acres of land. Express this area in square meters using scientific notation.
1 acre = 4.05×10^{-3} km^2.

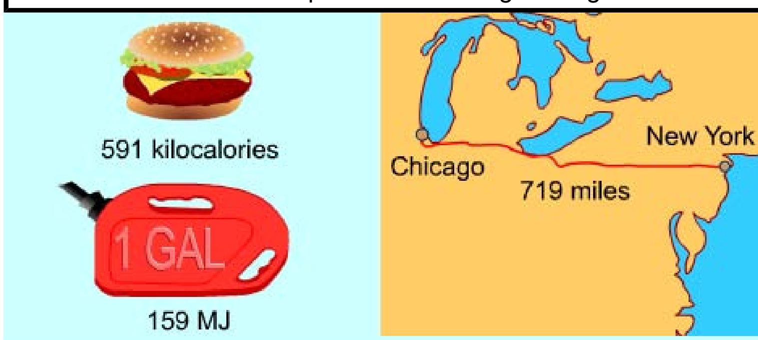
This interactive checkpoint allows you to practice scientific notation. The "e" symbol indicates that the leading value is multiplied by 10 raised to the power of the number that follows. For example, 1.35 e −6 is equal to 1.35×10^{-6}.

Here you will input the power into a separate box after the "e." In all other interactive checkpoints, the entire number (the leading value, e, and the power) is input into one box.

Answer:

$A =$ ______ e ______ m^2

1.16 - Interactive checkpoint: cheeseburgers or gasoline?



A double cheeseburger contains 591 kilocalories of energy. A gallon of gasoline contains 159 MJ of energy. A compact car makes 39.0 miles/gallon on the highway. Assuming that it could run as efficiently on fast food as on gasoline, find the number of cheeseburgers needed to propel the car from New York to Chicago (719 miles).

A kilocalorie is a unit used to measure food energy, and is often called a Calorie, spelled with a capital C. One calorie (small c) is equal to 4.19 J (joules).

Answer:

$N =$ ______ cheeseburgers

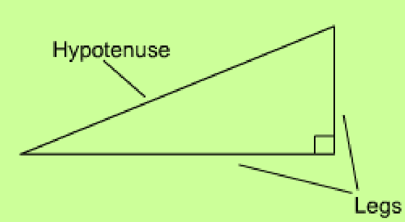
1.17 - Pythagorean theorem

As you proceed through your physics studies, you will find it necessary to understand the Pythagorean theorem, which is reviewed in this section. At the right, you see a right triangle (a triangle with a $90°$ angle). The Pythagorean theorem states that the square of the hypotenuse (the side opposite the right angle) equals the sum of the squares of the two legs. This equation is shown in Equation 1 to the right.

There are two specific right triangles that occur frequently in physics homework problems. In an *isosceles right triangle*, the two legs are the same length and the hypotenuse is the length of either leg times the square root of two. The angles of an isosceles right triangle are 45, 45 and 90 degrees, so it is also called a *45-45-90 triangle*.

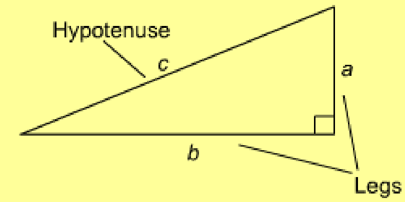
When the angles of the triangle measure 30, 60 and 90 degrees, it is called a *30-60-90 triangle*. The shorter leg, the one opposite the $30°$ angle, is one half the length of the hypotenuse. This relationship makes for an easy mathematical calculation and makes this triangle a favorite in homework problems.

concept 1



Pythagorean theorem
Relates hypotenuse to legs of right triangle

equation 1

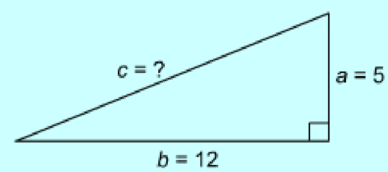


Pythagorean theorem

$$c^2 = a^2 + b^2$$

c = length of hypotenuse
a, b = lengths of legs

example 1



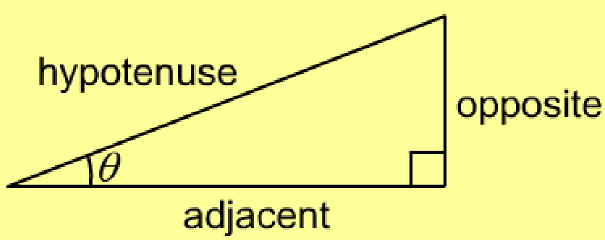
What is the length of the hypotenuse?

$c^2 = a^2 + b^2 = 5^2 + 12^2 = 25 + 144$

$c^2 = 169$

$c = \sqrt{169} = 13$

1.18 - Trigonometric functions



equation 1

Trigonometric functions

$\sin\theta$ = opposite / hypotenuse

$\cos\theta$ = adjacent / hypotenuse

$\tan\theta$ = opposite / adjacent

You will often encounter trigonometric functions in physics. You need to understand the basics of the sine, cosine and tangent, and their inverses: the arcsine, arccosine and arctangent.

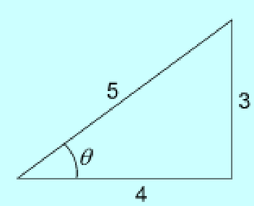
The illustration above depicts an angle and three sides of a triangle. The sine (sin) of the angle θ (the Greek letter *theta*, pronounced “thay-tuh”) equals the ratio of the side opposite the angle divided by the triangle's hypotenuse. “Opposite” means the leg across from the angle, as the diagram reflects.

The cosine (cos) of θ equals the ratio of the side of the triangle adjacent to the angle, divided by the hypotenuse. “Adjacent” means the leg that forms one side of the angle.

Finally, the tangent (tan) of θ equals the ratio of the opposite side divided by the adjacent side.

These three ratios are constant for a given angle in a right triangle, no matter what the size of the triangle. They are useful because you are often given information such as the length of the hypotenuse and the size of an angle, and then asked to calculate one

example 1



What is sin θ? cos θ? tan θ?

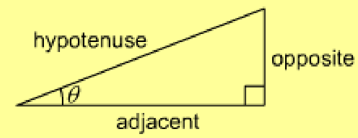
$\sin\theta$ = opposite / hypotenuse = 3/5 of the legs of the triangle. For instance, if asked to calculate the opposite leg, you would multiply the sine of the angle by the hypotenuse.

You may also be asked to use the arcsine, the arccosine or the arctangent. These are often written as $\sin^{-1}$, $\cos^{-1}$ and $\tan^{-1}$. These are not the reciprocals of the sine, cosine and tangent! Rather, they supply the size of the angle when the value of the trigonometric function is known. For example, since $\sin 30°$ equals 0.5, the arcsine of 0.5 (or $\sin^{-1}$ 0.5) equals $30°$. This is also often written as arcsin(0.5); arccos and arctan are the abbreviations for arccosine and arctangent.

In the old days, scientists consulted tables for these trigonometric values. Today, calculators and spreadsheets can calculate them for you.

$\cos\theta$ = adjacent / hypotenuse = 4/5
$\tan\theta$ = opposite / adjacent = 3/4

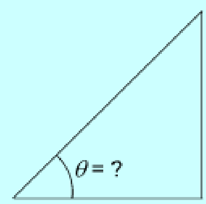
equation 2



Inverse trigonometric functions

θ = arcsin (opposite / hypotenuse)
θ = arccos (adjacent / hypotenuse)
θ = arctan (opposite / adjacent)
Often written: $\sin^{-1}$, $\cos^{-1}$, $\tan^{-1}$

example 2



If the tangent of θ is 1, what is θ?
θ = arctangent(1) = 45°

1.19 - Radians

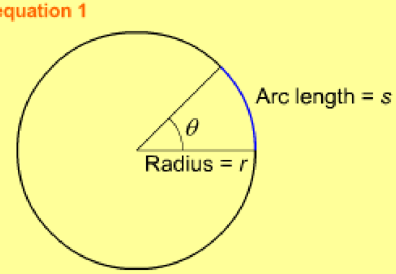
Radian measure: A measurement of angles based on a ratio of lengths.

Angles are often measured or specified in degrees, but another unit, the *radian*, is useful in many computations. The radian measure of an angle is the ratio of two lengths on a circle. The angle and lengths are perhaps most easily understood by looking at the diagram in Equation 1 on the right. The *arc length* is the length of the arc on the circumference cut off by the angle when it is placed at the circle's center. The other length is the radius of the circle. The radian measure of the angle equals the arc length divided by the radius.

A $360°$ angle equals 2π radians. Why is this so? The angle $360°$ describes an entire circle. The arc length in this case equals the circumference ($2\pi r$) of a circle divided by the radius r of the circle. The radius factor cancels out, leaving 2π as the result.

Radians are dimensionless numbers. Why? Since a radian is a ratio of two lengths, the length units cancel out. However, we follow a radian measure with "rad" so it is clear what is meant.

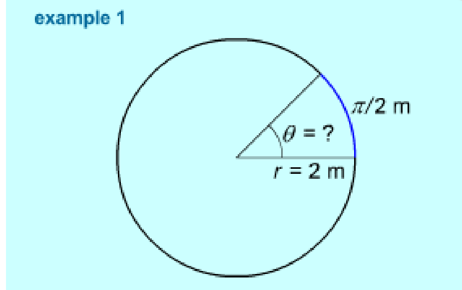
equation 1



Radian measure

Angle = arc length / radius = s/r
$360° = 2\pi$ rad

· Radians are dimensionless
· Units: radians (rad)



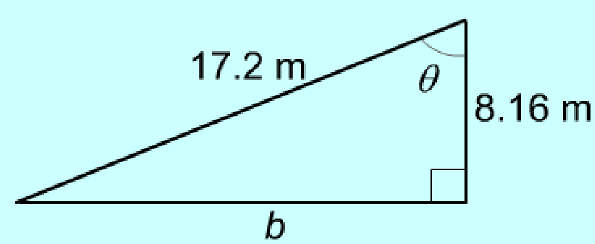
What is the angle's measure in radians?

Angle = arc length / radius

$\theta = (\pi/2 \text{ m})/(2 \text{ m})$

$\theta = \pi/4$ rad

1.20 - Sample problem: trigonometry



What is the length of side b of this triangle?

What is θ in degrees? in radians?

The lengths of two sides of a right triangle are shown.

Variables

short leg	a = 8.16 m
long leg	b
hypotenuse	c = 17.2 m
angle	θ

What is the strategy?

1. Use the Pythagorean theorem to calculate the length of the third side.
2. Calculate the cosine of θ and then use the arccosine function on a calculator (or consult a table) to determine θ in degrees.
3. Convert θ to radians.

Mathematics principles

$c^2 = a^2 + b^2$

$\cos\theta$ = adjacent/hypotenuse

$360° = 2\pi$ rad

Step-by-step solution

First we use the Pythagorean theorem to find the length of the longer leg.

Step	Reason
1. $c^2 = a^2 + b^2$	Pythagorean theorem
2. $(17.2 \text{ m})^2 = (8.16 \text{ m})^2 + b^2$	enter values
3. $b^2 = 17.2^2 - 8.16^2 = 229$	solve for b^2
4. b = 15.1 m	take square root

Now, we use the lengths of the sides to find the value of the cosine of θ, and then look up the arccosine of this ratio.

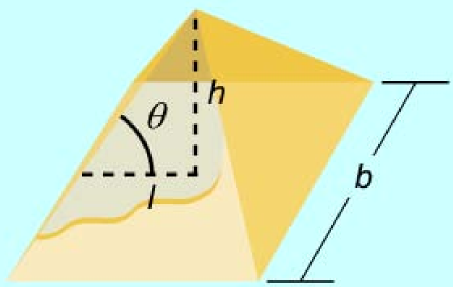
Step	Reason
5. $\cos\theta$ = adjacent/hypotenuse	definition of cosine
6. $\cos\theta = (8.16\text{ m})/(17.2\text{ m})$	enter values
7. $\cos\theta = 0.474$	divide
8. $\theta = \arccos(0.474)$	definition of arccosine
9. $\theta = 61.7°$	use calculator

Finally, we convert the angle from degrees to radians.

Step	Reason
10. $\theta = (61.7°)(2\pi\text{ rad}/360°)$	multiply by conversion factor
11. $\theta = 1.08$ rad	multiplication and division

There are other ways to solve this problem. For example, you could first find θ using the arccosine function, and then use the tangent ratio to find the length of the third side.

1.21 - Interactive checkpoint: trigonometry



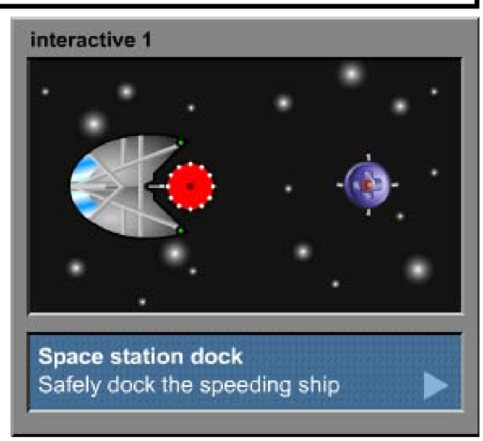
The great pyramid of Cheops has a square base with edges that are almost exactly 230 m long. The side faces of the pyramid make an angle of 51.8° with the ground.
The apex of the pyramid is directly above the center of the base. Find its height.

Answer:

h = ______ m

1.22 - Interactive summary problem: dock the shuttle

You are told the spacecraft to the right must accelerate at the rate of 2.15×10^5 miles per hour squared (mi/h^2) to dock with the space station. The simulation, however, only accepts values in meters per second squared (m/s^2), and it does not use scientific notation. You need to convert the units and then enter the value into the simulation to the right.



If you correctly convert the units and enter that value into the simulation, the spacecraft docks. If you enter an incorrect value, you may view a collision.

Click on the graphic to the right to start the simulation. Enter the converted value for the acceleration to the nearest 0.1 m/s^2. Press GO to make the spacecraft fly. Press RESET if you need to try again. (If you think a positive acceleration should cause the spacecraft to move faster, you will learn why this is not the case shortly.)

If you find this problem challenging, review the sections on converting units and scientific notation.

1.23 - Gotchas

The goal of “gotchas” is to help you avoid common errors. (Not that your teacher’s tests would ever try to make you commit any of these errors!)

Confusing weight and mass. You do not weigh 70 kilograms, or 80, or 60. However, those values could very well be your **mass**, which is an unchanging value that reflects your resistance to a change in motion.

Converting units with factors incorrectly oriented. There could probably be an essay written on this topic. *Make sure the units cancel!* is probably the best advice we can give. For example, if you are converting meters per second to miles per hour, begin by multiplying by a conversion fraction of 3600 seconds over one hour. This will cause the seconds to cancel and hours to be in the right place. (If this is not clear, write it down and strike out units. If it is still unclear, do some practice problems.) In any physics calculation, checking that the units on each side are consistent is a good technique.

Carelessly adding or subtracting numbers in scientific notation that have different exponents. The exponents must be the same before you can add or subtract the leading values.

1.24 - Summary

Scientists use the *Système International d'Unités*, also known as the metric system of measurement. Examples of metric units are meters, kilograms, and seconds.

In these systems, units that measure the same property, for example units for mass, are related to each other by powers of ten. Unit prefixes tell you how many powers of ten. For example, a kilogram is 1000 grams and a kilometer is 1000 meters, while a milligram is one one-thousandth of a gram, and a millimeter is one-thousandth of a meter.

Numbers may be expressed in scientific notation. Any number can be written as a number between 1 and 10, multiplied by a power of ten. For example, $875.6 = 8.756\times10^2$.

A standard is an agreed-on basis for establishing measurement units, like defining the kilogram as the mass of a certain platinum-iridium cylinder that is kept at the International Bureau of Weights and Measures, near Paris. A physical constant is an empirically measured value that does not change, such as the speed of light.

In the metric system, the basic unit of length is the meter; time is measured in seconds; and mass is measured in kilograms.

Sometimes a problem will require you to do unit conversion. Work in fractions so that you can cancel like units, and make sure that the units are of the same type (all are units of length, for instance).

When you need to do arithmetic using scientific notation, remember to deal with the leading values and the exponents separately. For multiplication, multiply the leading values and add the exponents. For division, divide the leading values and subtract the exponents. When adding or subtracting, first make sure the exponents are the same and then perform the operation on the leading values. In all cases, if the leading value of the result is not between one and 10, adjust the result. For example, 0.12×10^{-2} becomes 1.2×10^{-3}.

The Pythagorean theorem states that the square of the hypotenuse of a triangle is equal to the sum of the squares of the two legs.

$$c^2 = a^2 + b^2$$

Trigonometric functions, such as sine, cosine and tangent, relate the angles of a right triangle to the lengths of its sides.

Radians (rad) measure angles. The radian measure of an angle located at the center of a circle equals the arc length it cuts off on the circle, divided by the radius of the circle.

Equations

Prefixes

giga (G) = 10^9

mega (M) = 10^6

kilo (k) = 10^3

centi (c) = 10^{-2}

milli (m) = 10^{-3}

micro (μ) = 10^{-6}

nano (n) = 10^{-9}

Pythagorean Theorem

$c^2 = a^2 + b^2$

Trigonometric functions

$\sin\theta$ = opposite / hypotenuse

$\cos\theta$ = adjacent / hypotenuse

$\tan\theta$ = opposite / adjacent

Radian measure

Angle = arc length / radius = s/r

$360° = 2\pi$ rad

Chapter 1 Problems

Conceptual Problems

C.1 Why do scientists **not** define the standard for the second to be a fraction of a day?

C.2 What are the pitfalls in attempting to convert pounds into kilograms?

C.3 You take two measurements in a lab. The first measurement is in units called zorbs. The second measurement is in different units called zargs. Both zorbs and zargs have the same dimensions. Can you directly add these measurements together? If not, what additional information do you need in order to add them?

○ Yes ○ No

C.4 When multiplying by a conversion factor, how do you determine which unit belongs in the numerator and which in the denominator?

C.5 The following variables are commonly seen in equations. The name of the quantity represented by each variable, and its dimension(s), are also shown.
x distance (L)
t time (T)
m mass (M)
a acceleration (L/T^2)
v speed (L/T)
F force (ML/T^2)
Using the information above, check the boxes of the equations that are dimensionally correct.

Select all that apply.

☐ $F = ma$
☐ $v^2 = 2ax$
☐ $v = at^2$
☐ $F/v = m/t$

C.6 Give an example of a pair of different units with the same dimension.

C.7 You have verified that an equation is dimensionally correct. Is the equation necessarily true? Why?

○ Yes ○ No

C.8 Can a leg of a right triangle be longer than the hypotenuse? Why or why not?

○ Yes ○ No

C.9 Consider the sine and cosine functions. (a) As an angle increases from 0° to 90°, does the sine of the angle increase or decrease? (b) How about the cosine?

(a) ○ Increases ○ Decreases
(b) ○ Increases ○ Decreases

C.10 You walk along the edge of a large circular lawn. You walk clockwise from your starting location until you have moved an angle of π radians. What geometric shape is defined by the following three points: your starting position, your final position, and the center point of the lawn?

i. Isoceles triangle
ii. Right triangle
iii. Straight line

Section Problems

Section 2 - Prefixes

2.1 How many centimeters are there in a kilometer?

__________ cm

2.2 A carbon-carbon triple bond has a length of 120 picometers. What is its length in nanometers?

__________ nm

2.3 The 2000 Gross Domestic Product of the United States was $9,966 billion. If you were to regard the dollar as a scientific unit, how would you write this using the standard prefixes?

i.	9966	i.	pico dollars
ii.	9.966	ii.	milli
iii.	99.66	iii.	kilo
		iv.	tera

2.4 A gremlin is a tiny mythical creature often blamed for mechanical failures. Suppose for a moment that a shrewd physicist catches one such gremlin and makes it work for her instead of against her. She determines that the gremlin can produce 24 milliwatts of power. How many gremlins are required to produce 24,000 megawatts of power?

__________ gremlins

2.5 A drug company has just manufactured 50.0 kg of acetylsalicylic acid for use in aspirin tablets. If a single tablet contains 500 mg of the drug, how many tablets can the company make out of this batch?

__________ tablets

2.6 The diameter of an aluminum atom is about 0.24 nm, and the nuclear diameter is about 7.2 fm (femtometer = 10^{-15} meter). (a) If the atom's diameter were expanded to the length of an American football field (91.44 m) and the nuclear diameter expanded proportionally, what would be the nuclear diameter in meters? (b) Is the saying that "the atom is mostly empty space" confirmed by these figures?

(a) __________ m
(b) ○ Yes ○ No

Section 3 - Scientific notation

3.1 An electron can tunnel through an energy barrier with probability 0.0000000000375. (This is a concept used in quantum mechanics.) Express this probability in scientific notation.

____

3.2 An atom of uranium-235 has a mass 235.043924 amu (atomic mass units). Using scientific notation, state the mass in amu of ten million such atoms.

__________ amu

3.3 Sara has lived 18.0 years. How many seconds has she lived? Express the answer in scientific notation. Use 365.24 days per year for your calculations.

__________ s

3.4 Assume a typical hummingbird has a lifespan of 4.0 years and an average heart rate of 1300 beats per minute. (a) Calculate the number of times a hummingbird's heart beats in its life, and express it in scientific notation. Use 365.24 days per year. (b) A long-lived elephant lives for 61 years, and has an average heart rate of 25 beats per minute. Calculate the number of heartbeats in that elephant's lifetime in scientific notation.

(a) __________ beats
(b) __________ beats

3.5 A PC microprocessor runs at 2.40 GHz. A hertz (Hz) is a unit meaning "one cycle per second." A movie projector displays images at a rate of 24.0 Hz. What is the ratio of the microprocessor's rate in cycles per second to the update rate of a movie projector? Answer the question using scientific notation.

____

Section 5 - Length

5.1 Which of the following coins has a diameter of 17.91 mm: a nickel, a dime or a quarter?

i. Nickel
ii. Dime
iii. Quarter

Section 8 - Converting units

8.1 11.2 meters per second is how many miles per hour?

__________ mi/h

8.2 Freefall acceleration *g* is the acceleration due to gravity. It equals 9.80 meters per second squared near the Earth's surface. (a) What does it equal in feet per second squared? (b) In miles per second squared? (c) In miles per hour squared?

(a) __________ ft/s 2
(b) __________ mi/s 2
(c) __________ mi/h 2

8.3 A historian tells you that a cubit is an ancient unit of length equal to 0.457 meters. If you traveled 585 kilometers from Venice, Italy to Frankfurt, Germany, how many cubits did you cover?

__________ cubits

8.4 You are on the phone with a friend in Greece, who tells you that he has just caught a fish 65 cm long in the Mediterranean Sea. Assuming he is telling the truth, what is the length of the fish in inches?

__________ in

8.5 In an attempt to rid yourself of your little brother's unwanted attention, you tell him to count to a million and then come find you. If he were to start counting at a rate of one number per second, would he be finished in a year?

○ Yes ○ No

8.6 Light travels at 3.0×10^8 m/s in a vacuum. Find its speed in furlongs per fortnight. There are roughly 201 meters in a furlong, and a fortnight is equal to 14 days.

__________ furlongs/fortnight

8.7 Mercury orbits the Sun at a mean distance of 57,900,000 kilometers. (a) What is this distance in meters? Use scientific notation to express your answer and state it with three significant figures. (b) Pluto orbits at a mean distance of 5.91×10^{12} meters from the Sun. What is this distance in kilometers?

(a) __________ m
(b)
i. 5 910 000 km
ii. 5 910 000 000 km
iii. 591 000 000 000 km
iv. 5 910 000 000 000 km
v. 591 000 000 000 000 km

8.8 In 2003, Bill Gates was worth 40.7 billion dollars. (a) Express this figure in dollars in scientific notation. (b) Assume a dollar is worth 2,060 Italian lire. State Bill Gates's net worth in lire in scientific notation.

(a) __________ dollars
(b) __________ lire

8.9 The world's tallest man was Robert Pershing Wadlow, who was 8 feet, 11.1 inches tall. There are 2.54 centimeters in an inch and 12 inches in a foot. How tall was Robert in meters?

__________ m

8.10 Romans and Greeks used their stadium as a unit of measure. An estimate of the ancient Roman unit stadium is 185 m. A heroic epic describes a battle in which the Roman army marched 201 stadia to defend Rome from invading barbarians. In kilometers, how far did they march?

__________ km

8.11 You are carrying a 1.3 gallon jug full of cold water. Given that the water has a density of 1.0 grams per milliliter, and one gallon is equal to 3.8 liters, state the mass of the water in kilograms.

__________ kg

Section 9 - Interactive problem: converting units

9.1 Use the simulation in the interactive problem in this section to convert the speed of the car from miles per hour to meters per second.

__________ m/s

Section 10 - Dimensional analysis

10.1 The dimensions for force are the product of mass and length divided by time squared. Newton's second law states that force equals the product of mass and acceleration. What are the dimensions of acceleration?

○ T^2
○ T^2/L
○ L/T^2
○ L/T

10.2 The dimensions for force are the product of mass and length, divided by time squared. Newton's law of gravitation states that the gravitational force between two objects equals a constant, *G*, times the product of the mass of each object, divided by the square of the distance between them. What must the dimensions of the constant be?

- ○ L^3/MT^2
- ○ L^2/T^3
- ○ M^2T

10.3 The variables *x*, *v*, and *a* have dimensions L, L/T, and L/T^2, respectively. An equation relating these variables under particular conditions is $v^n = 2ax$, where *n* is a positive integer. What must *n* be?

____

10.4 The kinetic energy of an object is given by the equation $KE = (1/2)mv^2$, where *m* is mass and *v* is speed, with dimensions L/T. What are the dimensions of *KE*?

- ○ ML^2/T^2
- ○ L^2/T^2
- ○ LM/T

10.5 Here is a famous example worked out by the British physicist G.I. Taylor. In a nuclear explosion there is an essentially instantaneous release of energy (units: $kg{\cdot}m^2/s^2$). This produces a spherical shock wave. Taylor deduced that the radius of the sphere (units: meters) must depend only on the energy, the time *t* (units: seconds), and the undisturbed air density ρ (units: kg/m^3), with an undetermined constant. The constant is dimensionless (it has no units). Show how to combine the units of energy, density and time to obtain the units of length.

Section 11 - Multiplying numbers in scientific notation

11.1 Multiply $3.65{\times}10^{23}$ by $4.12{\times}10^{154}$ by $1.11{\times}10^{-11}$ and express the answer in scientific notation.

____

11.2 Evaluate $(5.7{\times}10^6\ kg) \times (6.3{\times}10^{-2}\ m/s^2)$ and express the answer in scientific notation.

__________ kg 'd7 m/s 2

11.3 What is $(9.80{\times}10^8) \times (-2.02{\times}10^{-4})$? Express the answer in scientific notation.

____

11.4 There are many estimates of the number of stars in the universe. One estimate is that there are 10^{10} stars in the Milky Way galaxy, and that there are 10^{10} galaxies in the universe. Assuming that the number of stars in the Milky Way is the average number of stars in a galaxy, estimate how many stars there are in the universe. It perhaps goes without saying: Express your answer in scientific notation.

____

Section 12 - Dividing numbers in scientific notation

12.1 According to the CIA, about $2.69{\times}10^6$ people were living in Jamaica in 2003. The organization estimated the gross domestic product (GDP) of Jamaica as $10.1 billion at about this time. (a) Express the GDP in scientific notation. (b) What was the GDP per person? Express the number in scientific notation.

(a) __________ dollars
(b) __________ dollars per person

12.2 Evaluate $(4.9{\times}10^{-8}) / (7.0{\times}10^{-3})$. Express the answer in scientific notation.

____

12.3 Newton's second law states that the net force equals the product of mass and acceleration. A boat's mass of is $9.6{\times}10^5$ kg and it experiences a net force of $1.5{\times}10^4\ kg{\cdot}m/s^2$. State its acceleration.

__________ m/s 2

12.4 Average speed equals distance divided by time. On a journey to the Alpha Centauri star system, a space probe travels a distance of $4.12{\times}10^{16}$ meters in $1.73{\times}10^{10}$ seconds. (a) Find its average speed in meters per second. (b) How long did the journey take in years? Use 365.24 days per year.

(a) __________ m/s
(b) __________ years

12.5 An estimate for the number of cells in a human body is 50 million million (yes, that is two millions in a row). What percentage of your body mass does one average cell constitute? Express your answer in scientific notation.

__________ %

12.6 The force of gravitation can be calculated with the equation $F = Gm_1m_2/r^2$ where $G = 6.67\times10^{-11}$ N·m²/kg². Given that $m_1 = 1.50\times10^{15}$ kg, $m_2 = 2.40\times10^{15}$ kg, and $r = 6.20\times10^6$ m, find F.

__________ N

Section 13 - Adding and subtracting numbers in scientific notation

13.1 A 3.70×10^6 kg piece splits off an iceberg of mass of 5.96×10^7 kg. Calculate the mass of the remaining iceberg and express the answer in scientific notation.

__________ kg

13.2 The answer to each of the following computations in scientific notation is incorrect. Give the correct answer for each part, and explain the error made in the computation shown.

(a) $(6.0\times10^5) / (3.0\times10^3) = 2.0\times10^8$
(b) $(3.7\times10^8) / (7.6\times10^4) = 0.5\times10^4$
(c) $(4.8\times10^{13}) + (4.8\times10^{12}) = 9.6\times10^{13}$
(d) $(2.3\times10^3) \times (1.4\times10^3) = 3.2\times10^9$

(a) __________
(b) __________
(c) __________
(d) __________

13.3 The silicon chip in a microelectronic component has a mass of 1.37×10^{-3} kg. To complete the wiring, 6.79×10^{-5} kg of solder is added to the chip. Compute the final mass of the circuit.

__________ kg

13.4 You have $\$1.14\times10^4$ in your checking account but must pay $\$3.30\times10^3$ in tuition. What is the balance in your checking account after you pay your tuition? Express the answer in scientific notation.

__________ dollars

13.5 A small metal chamber contains 5.632×10^{-9} mol of oxygen gas at very low pressure. An additional 4.379×10^{-8} mol of oxygen gas are added. Find the total amount of gas in the chamber.

__________ mol

13.6 An auto transport truck has a mass of 2.50×10^4 kg. A car has a mass of 1.31×10^3 kg. The car is loaded onto the truck. What is their total mass?

__________ kg

13.7 Harry carries a backpack that weighs 8.03×10^3 g. He steps on a scale wearing his backpack and the scale reads 7.62×10^4 g. What is Harry's mass in grams?

__________ g

Section 17 - Pythagorean theorem

17.1 A crew of piano movers uses a 6.5-foot ramp to move a 990 lb Steinway concert grand up onto a stage that is 1.8 feet higher than the floor. How far away from the base of the stage should they set the end of the ramp so that the other end exactly reaches the stage?

__________ ft

17.2 Suzy is holding her kite on a string 25.0 m long when the kite hits the top of a flagpole, which is 15.3 m higher than her hands. Assuming that the string is taut and forms a straight line, what is the horizontal distance from her hands to the flagpole?

__________ m

17.3 A right triangle has a hypotenuse of length 17 cm and a leg of length 15 cm. What is the length of the other leg?

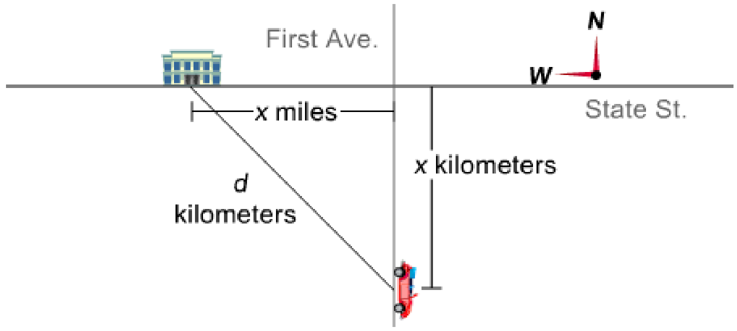
__________ cm

17.4 A Pythagorean triple is a set of three integers (a,b,c) that could form three sides of a right triangle. (3,4,5) and (5,12,13) are two examples. There exists a Pythagorean triple of the form (7,n,n+1). Find n.

____

17.5 A hapless motorist is trying to find his friend's house, which is located 2.00 miles west of the intersection of State Street and First Avenue. He sets out from the intersection, and drives the same number of kilometers due south instead. How far in kilometers is he from his intended destination? The sketch is not to scale.

__________ km



Section 18 - Trigonometric functions

18.1 You want to estimate the height of the Empire State Building. You start at its base and walk 15 m away. Then you approximate the angle from the ground at that point to the top to be 88 degrees. How tall do you estimate the Empire State Bulding to be?

__________ m

18.2 For an angle θ, write tan θ in terms of sin θ and cos θ.

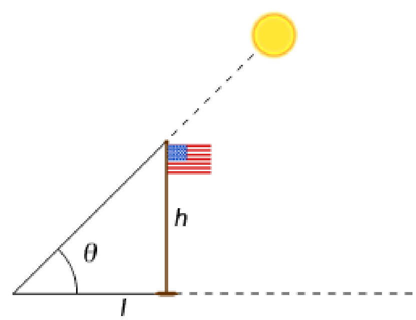
- ○ tan θ = cos θ / sin θ
- ○ tan θ = sin θ / cos θ
- ○ tan θ = (sin θ)(cos θ)

18.3 A window washer is climbing up a ladder to wash a window. The end of the 10.5-foot ladder exactly touches the windowsill and the ladder makes a 70.0° angle with the ground. How far off the ground is the windowsill?

__________ ft

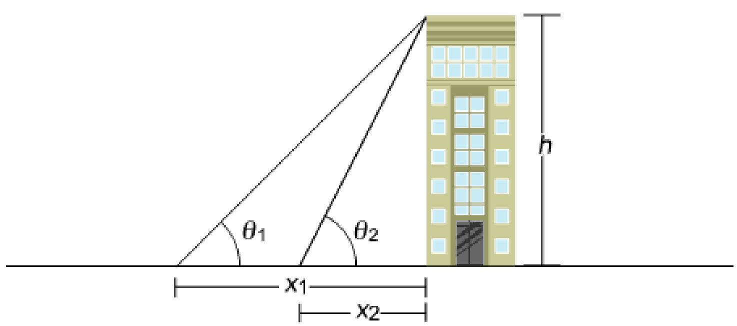
18.4 The 28 ft flagpole outside Martin Luther King, Jr. Elementary School casts a 42 ft shadow. In degrees, at what angle above the horizon is the Sun? The diagram may not be to scale.

__________ °



18.5 A superhero is running toward a skyscraper to stop a villain from getting away. As he is moving, he looks up at the top of the building and his line of sight makes a 33.0° angle with the ground. After having moved 250 m, he looks at the top of the building again, and notices that his line of sight makes a 52.0° angle with the horizontal. How tall is the skyscraper?

__________ m



Section 19 - Radians

19.1 Find the radian measure of one of the angles of a regular pentagon. (The angles are all the same in a regular pentagon.) In degrees, the formula for the sum of the interior angles of an n-sided polygon is $(n - 2) \cdot 180°$.

__________ radians

19.2 A right triangle has sides of length 1, 2 and the square root of 3. In radians, what is its smallest angle?

- ○ $\pi/3$ radians
- ○ $\pi/6$ radians
- ○ $\pi/12$ radians

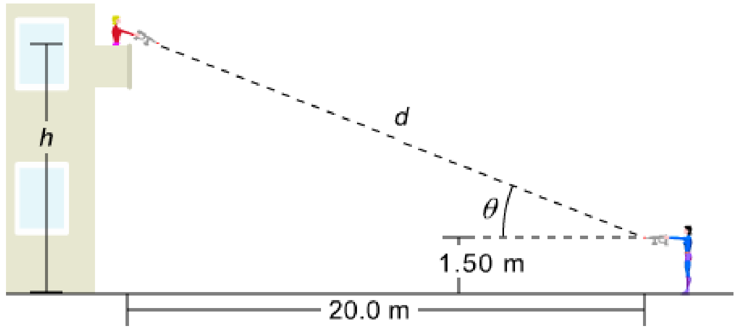
Section 22 - Interactive summary problem: dock the shuttle

22.1 Use the simulation in the interactive problem in this section to convert the acceleration from mi/hr^2 to m/s^2.

_________ m/s 2

Additional Problems

A.1 You are playing a game of laser tag with a friend who is on a balcony 12 meters off the ground. You are standing so your laser is 20.0 meters away from your friend (measured horizontally), and 1.50 meters off the ground. (a) At what angle from the horizontal should you aim the laser to hit your friend? Express your answer in degrees. (b) What is this angle in radians? (c) What distance does the laser light travel?



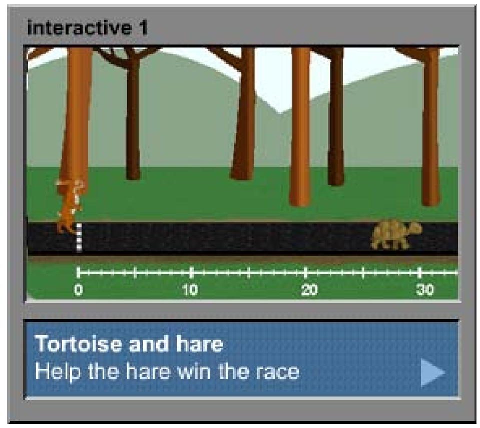
(a) _________ °
(b) _________ rad
(c) _________ m

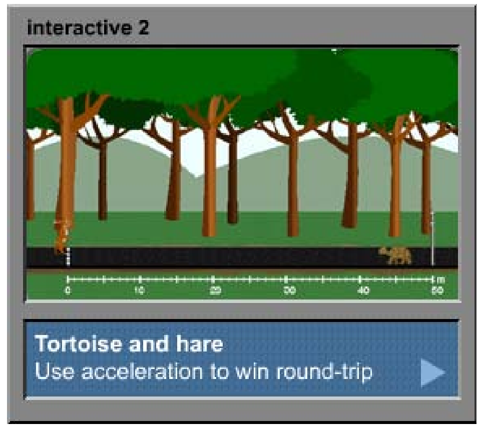
A.2 A freight version of Boeing's 747 airliner has 2.7×10^4 cubic feet of cargo space. A CD in its jewel case has a volume of 10.8 cubic inches. Assume that each CD holds 650 megabytes of information. (a) If you packed the aircraft with as many CDs as it could hold, how much information could you store in megabytes? (b) A "56K" modem typically transmits information at 45 kilobits per second. In computer terminology, a kilobit is 1024 bits, and a megabyte is 1,048,576 bytes. One byte equals eight bits of information. How long would it take to transmit the amount of information contained in one 747 planeful of CDs through a modem? (c) A T3 high-speed internet connection transmits information at 45.0 megabits per second. As with bytes and bits, one megabyte equals eight megabits of information. Assume that it takes 12 hours to load, fly and unload the 747. How many T3 connections would it take to transmit this information in 12 hours?

(a) _________ megabytes
(b) _________ seconds
(c) _________ connections

chapter 2
Motion in One Dimension

2.0 - Introduction





Objects move: Balls bounce, cars speed, and spaceships accelerate. We are so familiar with the concept of motion that we use sophisticated physics terms in everyday language. For example, we might say that a project has reached “escape velocity” or, if it is going less well, that it is in “free fall.”

In this chapter, you will learn more about motion, a field of study called *kinematics*. You will become familiar with concepts such as velocity, acceleration and displacement. For now, the focus is on how things move, not what causes them to move. Later, you will study *dynamics*, which centers on forces and how they affect motion. Dynamics and kinematics make up *mechanics*, the study of force and motion.

Two key concepts in this chapter are velocity and acceleration. Velocity is how fast something is moving (its speed) **and** in what direction it is moving. Acceleration is the rate of change in velocity. In this chapter, you will have many opportunities to learn about velocity and acceleration and how they relate. To get a feel for these concepts, you can experiment by using the two simulations on the right. These simulations are versions of the tortoise and hare race. In this classic parable, the steady tortoise always wins the race. With your help, though, the hare stands a chance. (After all, this is your physics course, not your literature course.)

In the first simulation, the tortoise has a head start and moves at a constant velocity of three meters per second to the right. The hare is initially stationary; it has zero velocity. You set its acceleration − in other words, how much its velocity changes each second. The acceleration you set is constant throughout the race. Can you set the acceleration so that the hare crosses the finish line first and wins the race? To try, click on Interactive 1, enter an acceleration value in the entry box in the simulation, and press GO to see what happens. Press RESET if you want to try again. Try acceleration values up to 10 meters per second squared. (At this acceleration, the velocity increases by 10 meters per second every second. Values larger than this will cause the action to occur so rapidly that the hare may quickly disappear off the screen.)

It does not really matter if you can cause the hare to beat this rather fast-moving tortoise. However, we do want you to try a few different rates of acceleration and see how they affect the hare’s velocity. Nothing particularly tricky is occurring here; you are simply observing two basic properties of motion: velocity and acceleration.

In the second simulation, the race is a round trip. To win the race, a contestant needs to go around the post on the right and then return to the starting line. The tortoise has been given a head start in this race. When you start the simulation, the tortoise has already rounded the post and is moving at a constant velocity on the homestretch back to the finish line.

In this simulation, when you press GO the hare starts off moving quickly to the right. Again, you supply a value for its acceleration. The challenge is to supply a value for the hare's acceleration so that it turns around at the post and races back to beat the tortoise. (Hint: Think negative! Acceleration can be either positive or negative.)

Again, it does not matter if you win; we want you to notice how acceleration affects velocity. Does the hare's velocity ever become zero? Negative? To answer these questions, click on Interactive 2, enter the acceleration value for the hare in the gauge, press GO to see what happens, and RESET to try again. You can also use PAUSE to stop the action and see the velocity at any instant. Press PAUSE again to restart the race.

We have given you a fair number of concepts in this introduction. These fundamentals are the foundation of the study of motion, and you will learn much more about them shortly.

2.1 - Position

Position: The location of an object; in physics, typically specified with graph coordinates.

Position tells you location.

There are many ways to describe location: Beverly Hills 90210; “...a galaxy far, far away”; “as far away from you as possible.” Each works in its own context.

Physicists often use numbers and graphs instead of words and phrases. Numbers and graphs enable them (and you) to analyze motion with precision and consistency. In this chapter, we will analyze objects that move in one dimension along a line, like a train moving along a flat, straight section of track.

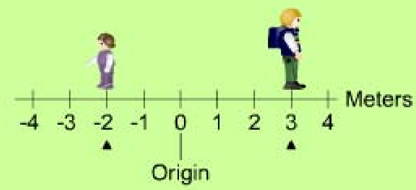
To begin, we measure position along a *number line*. Two toy figures and a number line are shown in the illustrations to the right. As you can see, the zero point is called the *origin*. Positive numbers are on the right and negative numbers are on the left.

By convention, we draw number lines from left to right. The number line could reflect an object's position in east and west directions, or north and south, or up and down; the important idea is that we can specify positions by referring to points on a line.

When an object moves in one dimension, you can specify its position by its location on the number line. The variable x specifies that position. For example, as shown in the illustration for Equation 1, the hiker stands at position $x = 3.0$ meters and the toddler is at position $x = -2.0$ meters.

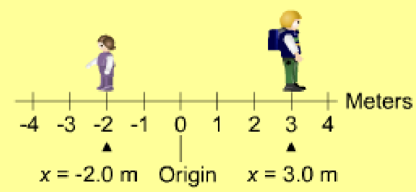
Later, you will study objects that move in multiple dimensions. For example, a basketball free throw will initially travel both up and forward. For now, though, we will consider objects that move in one dimension.

concept 1



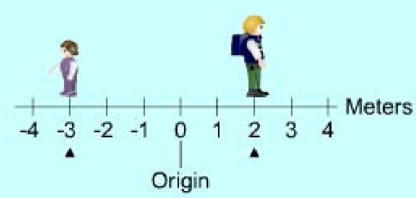
Position
Location of an object
Relative to origin

equation 1



Position
x represents position
Units: meters (m)

example 1



What are the positions of the figures?
Hiker: $x = 2.0$ m
Toddler: $x = -3.0$ m

2.2 - Displacement

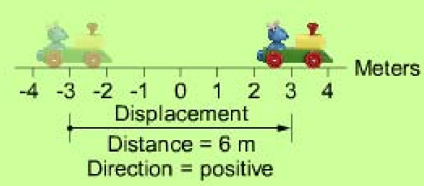
Displacement: The direction and distance of the shortest path between an initial and final position.

You use the concept of distance every day. For example, you are told a home run travels 400 ft (122 meters) or you run the metric mile (1.5 km) in track (or happily watch others run a metric mile).

Displacement adds the concept of direction to distance. For example, you go approximately 954 mi (1540 km) **south** when you travel from Seattle to Los Angeles; the summit of Mount Everest is 29,035 ft (8849.9 meters) **above** sea level. (You may have noticed we are using both metric and English units. We will do this only for the first part of this chapter, with the thought that this may prove helpful if you are familiarizing yourself with the metric system.)

Sometimes just distance matters. If you want to be a million miles away from your

concept 1



Displacement
Distance and direction
Measures *net* change in position

younger brother, it does not matter whether that's east, north, west or south. The distance is called the magnitude − the amount − of the displacement.

Direction, however, can matter. If you walk 10 blocks north of your home, you are at a different location than if you walk 10 blocks south. In physics, direction often matters. For example, to get a ball to the ground from the top of a tall building, you can simply drop the ball. Throwing the ball back up requires a very strong arm. Both the direction and distance of the ball's movement matters.

The definition of displacement is precise: the direction and length of the **shortest** path from the **initial** to the **final** position of an object's motion. As you may recall from your mathematics courses, the shortest path between two points is a straight line. Physicists use arrows to indicate the direction of displacement. In the illustrations to the right, the arrow points in the direction of the mouse's displacement.

Physicists use the Greek letter Δ (delta) to indicate a change or difference. A change in position is displacement, and since x represents position, we write Δx to indicate displacement. You see this notation, and the equation for calculating displacement, to the right. In the equation, x_f represents the final position (the subscript f stands for final) and x_i represents the initial position (the subscript i stands for initial).

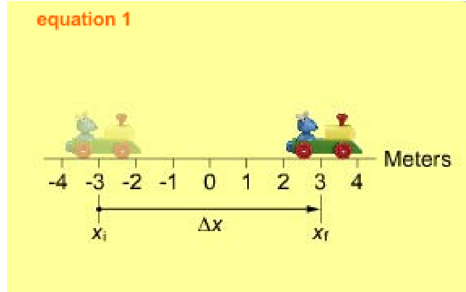
Displacement is a vector. A vector is a quantity that must be stated in terms of its direction and its magnitude. Magnitude means the size or amount. "Move five meters to the right" is a description of a vector. Scalars, on the other hand, are quantities that are stated solely in terms of magnitude, like "a dozen eggs." There is no direction for a quantity of eggs, just an amount.

In one dimension, a positive or negative sign is enough to specify a direction. As mentioned, numbers to the right of the origin are positive, and those to the left are negative. This means displacement to the right is positive, and to the left it is negative. For instance, you can see in Example 1 that the mouse's car starts at the position +3.0 meters and moves to the left to the position −1.0 meters. (We measure the position at the middle of the car.) Since it moves to the left 4.0 meters, its displacement is −4.0 meters.

Displacement measures the distance solely between the beginning and end of motion. We can use dance to illustrate this point. Let's say you are dancing and you take three steps forward and two steps back. Although you moved a total of five steps, your displacement after this maneuver is one step forward.

It would be better to use signs to describe the dance directions, so we could describe forward as "positive" and backwards as "negative." Three steps forward and two steps back yield a displacement of positive one step.

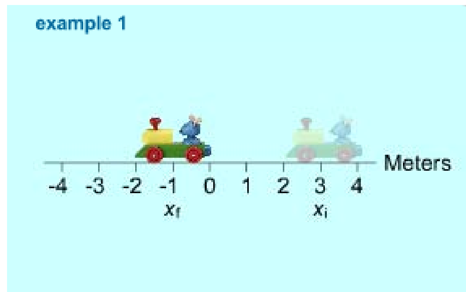
Since displacement is in part a measure of distance, it is measured with units of length. Meters are the SI unit for displacement.



Displacement

$$\Delta x = x_f - x_i$$

Δx = displacement
x_f = final position
x_i = initial position
Units: meters (m)



What is the mouse car's displacement?

$\Delta x = x_f - x_i$

$\Delta x = -1.0 \text{ m} - 3.0 \text{ m}$

$\Delta x = -4.0 \text{ m}$

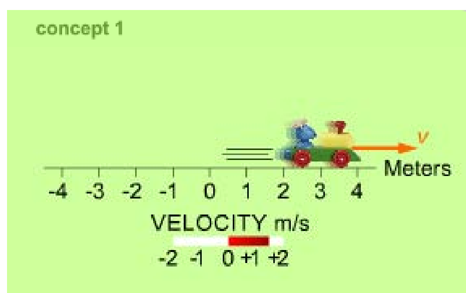
2.3 - Velocity

Velocity: Speed and direction.

You are familiar with the concept of speed. It tells you how fast something is going: 55 miles per hour (mi/h) is an example of speed. The speedometer in a car measures speed but does not indicate direction.

When you need to know both speed and direction, you use velocity. Velocity is a vector. It is the measure of how fast **and** in which direction the motion is occurring. It is represented by v. In this section, we focus on average velocity, which is represented by v with a bar over it, as shown in Equation 1.

A police officer uses the concepts of both speed and velocity in her work. She might issue a ticket to a motorist for driving 36 mi/h (58 km/h) in a school zone; in this case, speed matters but direction is irrelevant. In another situation, she might be told that a suspect is fleeing **north** on I-405 at 90 mi/h (149 km/h); now velocity is important because it tells her both how fast and in what direction.



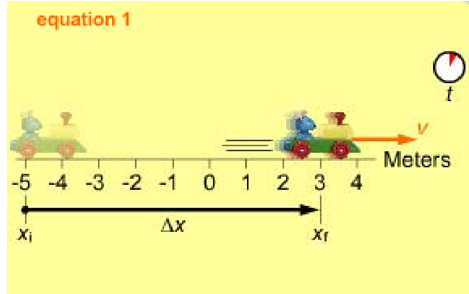
Velocity
Speed and direction

To calculate an object's average velocity, divide its displacement by the time it takes to move that displacement. This time is called the elapsed time, and is represented by Δt. The direction for velocity is the same as for the displacement.

For instance, let's say a car moves positive 50 mi (80 km) between the hours of 1 P.M. and 3 P.M. Its displacement is positive 50 mi, and two hours elapse as it moves that distance. The car's average velocity equals +50 miles divided by two hours, or +25 mi/h (+40 km/h). Note that the direction is positive because the displacement was positive. If the displacement were negative, then the velocity would also be negative.

At this point in the discussion, we are intentionally ignoring any variations in the car's velocity. Perhaps the car moves at constant speed, or maybe it moves faster at certain times and then slower at others. All we can conclude from the information above is that the car's **average** velocity is +25 mi/h.

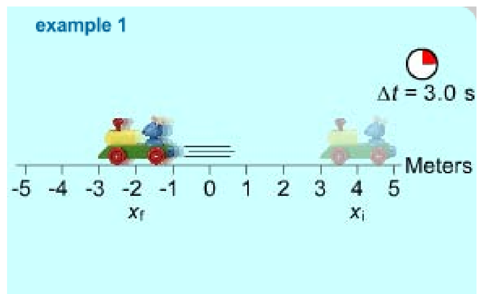
Velocity has the dimensions of length divided by time; the units are meters per second (m/s).



Velocity

$$\bar{v} = \Delta x / \Delta t$$

$\bar{v}$ = (average) velocity
Δx = displacement
Δt = elapsed time
Units: meters/second (m/s)



What is the mouse's velocity?

$\bar{v} = \Delta x / \Delta t$

$\bar{v} = (-2.0\ \text{m} - 4.0\ \text{m}) / 3.0\ \text{s}$

$\bar{v} = (-6.0\ \text{m}) / (3.0\ \text{s})$

$\bar{v} = -2.0\ \text{m/s}$

2.4 - Average velocity

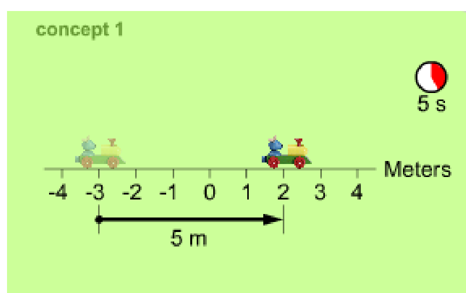
Average velocity: Displacement divided by elapsed time.

Average velocity equals displacement divided by the time it takes for the displacement to occur.

For example, if it takes you two hours to move positive 100 miles (160 kilometers), your average velocity is +50 mi/h (80 km/h). Perhaps you drive a car at a constant velocity. Perhaps you drive really fast, slow down for rush-hour traffic, drive fast again, get pulled over for a ticket, and then drive at a moderate speed. In either case, because your displacement is 100 mi and the elapsed time is two hours, your average velocity is +50 mi/h.

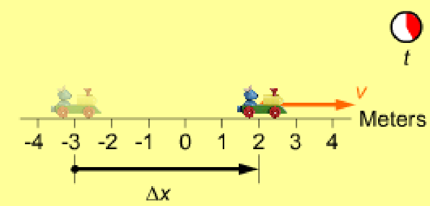
Since the average velocity of an object is calculated from its displacement, you need to be able to state its initial and final positions. In Example 1 on the right, you are shown the positions of three towns and asked to calculate the average velocity of a trip. You must calculate the displacement from the initial to final position to determine the average velocity.

A classic physics problem tempts you to err in calculating average velocity. The problem runs like this: "A hiker walks one mile at two miles per hour, and the next mile at four miles per hour. What is the hiker's average velocity?" If you average two and four and answer that the average velocity is three mi/h, you will have erred. To answer the problem, you must first calculate the elapsed time. You cannot simply average the two velocities. It takes the hiker 1/2 an hour to cover the first mile, but only 1/4 an hour to walk the second mile, for a total elapsed time of 3/4 of an hour. The average velocity equals two miles divided by 3/4 of an hour, which is a little less than three miles per hour.



Average velocity
Displacement divided by elapsed time

equation 1

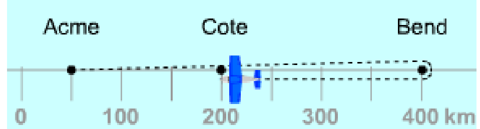


Average velocity

$$\bar{v} = \Delta x / \Delta t$$

$\bar{v}$ = average velocity
Δx = displacement
Δt = elapsed time

example 1



A plane flies Acme to Bend in 2.0 hrs, then straight back to Cote in 1.0 hr. What is its average velocity for the trip in km/h?

$$\bar{v} = \Delta x / \Delta t = (x_f - x_i) / \Delta t$$

$$\bar{v} = (200\text{ km} - 50\text{ km})/3.0\text{ h}$$

$$\bar{v} = 50\text{ km/h}$$

2.5 - Instantaneous velocity

Instantaneous velocity: Velocity at a specific moment.

Objects can speed up or slow down, or they can change direction. In other words, their velocity can change. For example, if you drop an egg off a 40-story building, the egg's velocity will change: It will move faster as it falls. Someone on the building's 39th floor would see it pass by with a different velocity than would someone on the 30th.

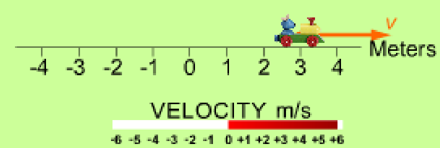
When we use the word "instantaneous," we describe an object's velocity at a particular instant. In Concept 1, you see a snapshot of a toy mouse car at an instant when it has a velocity of positive six meters per second.

The fable of the tortoise and the hare provides a classic example of instantaneous versus average velocity. As you may recall, the hare seemed faster because it could achieve a greater instantaneous velocity than could the tortoise. But the hare's long naps meant that its average velocity was less than that of the tortoise, so the tortoise won the race.

When the average velocity of an object is measured over a very short elapsed time, the result is close to the instantaneous velocity. The shorter the elapsed time, the closer the average and instantaneous velocities. Imagine the egg falling past the 39th floor window in the example we mentioned earlier, and let's say you wanted to determine its instantaneous velocity at the midpoint of the window.

You could use a stopwatch to time how long it takes the egg to travel from the top to the bottom of the window. If you then divided the height of the window by the elapsed time,

concept 1



Instantaneous velocity
Velocity at a specific moment

equation 1

Instantaneous velocity

$$v = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$$

v = instantaneous velocity

the result would be close to the instantaneous velocity. However, if you measured the time for the egg to fall from 10 centimeters above the window's midpoint to 10 centimeters below, and used that displacement and elapsed time, the result would be even closer to the instantaneous velocity at the window's midpoint. As you repeated this process "to the limit" − measuring shorter and shorter distances and elapsed times (perhaps using motion sensors to provide precise values) − you would get values closer and closer to the instantaneous velocity.

Δx = displacement
Δt = elapsed time

To describe instantaneous velocity mathematically, we use the terminology shown in Equation 1. The arrow and the word "lim" mean the limit as Δt approaches zero. The limit is the value approached by the calculation as it is performed for smaller and smaller intervals of time.

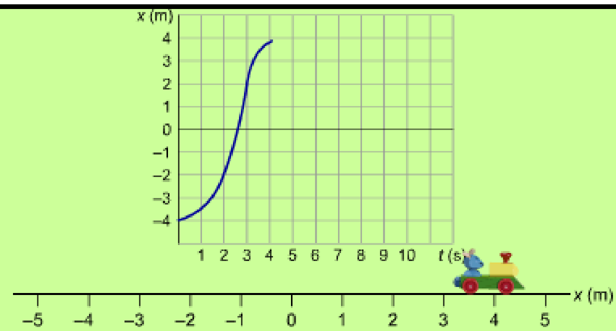
To give you a sense of velocity and how it changes, let's again use the example of the egg. We calculate the velocity at various times using an equation you may have not yet encountered, so we will just tell you the results. Let's assume each floor of the building is four meters (13 ft) high and that the egg is being dropped in a vacuum, so we do not have to worry about air resistance slowing it down.

One second after being dropped, the egg will be traveling at 9.8 meters per second; at three seconds, it will be traveling at 29 m/s; at five seconds, 49 m/s (or 32 ft/s, 96 ft/s and 160 ft/s, respectively.)

After seven seconds, the egg has an instantaneous velocity of 0 m/s. Why? The egg hit the ground at about 5.7 seconds and therefore is not moving. (We assume the egg does not rebound, which is a reasonable assumption with an egg.)

Physicists usually mean "instantaneous velocity" when they say "velocity" because instantaneous velocity is often more useful than average velocity. Typically, this is expressed in statements like "the velocity when the elapsed time equals three seconds."

2.6 - Position-time graph and velocity



concept 1

Position-time graph
Shows position of object over time
Steeper graph = greater speed

A graph of an object's position over time is a useful tool for analyzing motion. You see such a position-time graph above. Values on the vertical axis represent the mouse car's position, and time is plotted on the horizontal axis. You can see from the graph that the mouse car started at position $x = -4$ m, then moved to the position $x = +4$ m at about $t = 4.5$ s, stayed there for a couple of seconds, and then reached the position $x = -2$ m again after a total of 12 seconds of motion.

Where the graph is horizontal, as at point B, it indicates the mouse's position is not changing, which is to say the mouse is not moving. Where the graph is steep, position is changing rapidly with respect to time and the mouse is moving quickly.

Displacement and velocity are mathematically related, and a position-time graph can be used to find the average or instantaneous velocity of an object. The slope of a straight line between any two points of the graph is the object's **average** velocity between them.

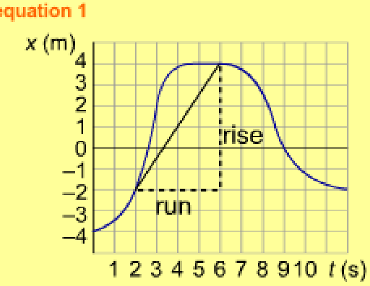
Why is the average velocity the same as this slope? The slope of a line is calculated by dividing the change in the vertical direction by the change in the horizontal direction, "the rise over the run." In a position-time graph, the vertical values are the x positions and the horizontal values tell the time. The slope of the line is the change in position, which is displacement, divided by the change in time, which is the elapsed time. This is the definition of average velocity: displacement divided by elapsed time.

You see this relationship stated and illustrated in Equation 1. Since the slope of the line shown in this illustration is positive, the average velocity between the two points on the line is positive. Since the mouse moves to the right between these points, its displacement is positive, which confirms that its average velocity is positive as well.

The slope of the tangent line for any point on a straight-line segment of a position-time graph is constant. When the slope is constant, the velocity is constant. An example of constant velocity is the horizontal section of the graph that includes the point B in the illustration above.

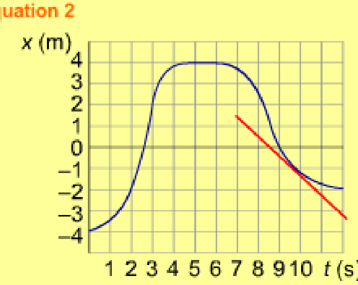
The slope of a tangent line at different points on a curve is **not** constant. The slope at a single point on a curve is determined by the slope of the tangent line to the curve at that point. You see a tangent line illustrated in Equation 2. The slope as measured by the tangent line equals the **instantaneous** velocity at the point. The slope of the tangent line in Equation 2 is negative, so the velocity there is negative. At that point, the mouse is moving from right to left. The negative displacement over a short time interval confirms that its velocity is negative.

equation 1



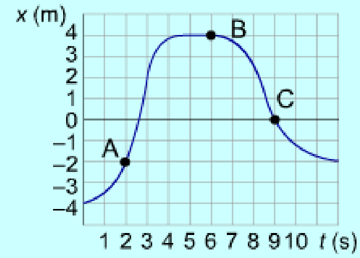
Average velocity
Slope of line between two points
$\bar{v} = \text{rise/run} = \Delta x / \Delta t$

equation 2



Instantaneous velocity
Slope of tangent line at point

example 1



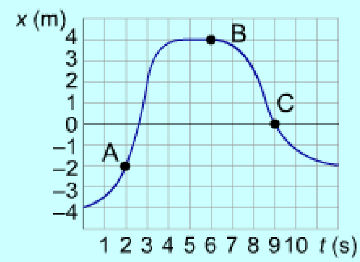
What is the average velocity between points A and B?

$\bar{v} = \text{rise/run}$

$$\bar{v} = \frac{4.0 \text{ m} - (-2.0 \text{ m})}{6.0 \text{ s} - 2.0 \text{ s}}$$

$\bar{v} = 1.5 \text{ m/s}$

example 2



Consider the points A, B and C. Where is the instantaneous velocity zero? Where is it positive? Where is it negative?

Zero at B

Positive at A

Negative at C

2.7 - Interactive problem: draw a position-time graph

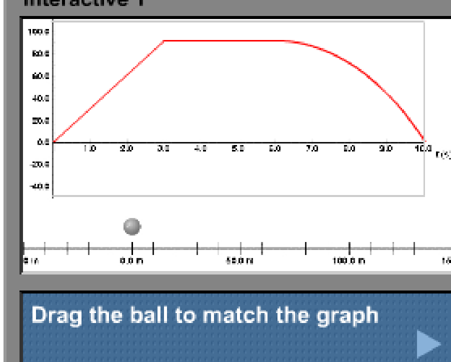
In this section, you are challenged to match a pre-drawn position-time graph by moving a ball along a number line. As you drag the ball, its position at each instant will be graphed. Your challenge is to get as close as you can to the target graph.

When you open the interactive simulation on the right, you will see a graph and a coordinate system with x positions on the vertical axis and time on the horizontal axis. Below the graph is a ball on a number line. Examine the graph and decide how you will move the ball over the 10 seconds to best match the target graph. You may find it helpful to think about the velocity described by the target graph. Where is it increasing? decreasing? zero? If you are not sure, review the section on position-time graphs and velocity.

You can choose to display a graph of the velocity of the motion of the ball as described by the target graph by clicking a checkbox. We encourage you to think first about what the velocity will be and use this checkbox to confirm your hypothesis.

Create your graph by dragging the ball and watching the graph of its motion. You can press RESET and try again as often as you like.

interactive 1



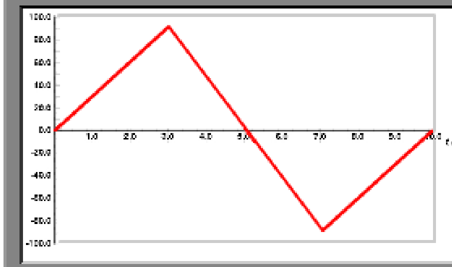
2.8 - Interactive problem: match a graph using velocity

In the interactive simulation on the right, you see a position-time graph of a ball that moves horizontally. The target graph has three straight-line segments: from 0 to 3 seconds, 3 to 7 seconds, and 7 to 10 seconds. You set the velocity for three time intervals.

Your challenge is to set the velocity for each interval to match the target graph. The velocity is constant for each of the intervals. You can calculate each velocity from the slope of the appropriate line segment. Review the section on position-time graphs and velocity if you are not sure how to do this.

Enter the velocity for each segment to the nearest meter per second and press GO to start the ball moving. Press RESET to start over.

interactive 1



Set three velocities to match the graph

2.9 - Velocity graph and displacement

If you graph an object's velocity versus time, the area between the graph and the horizontal axis equals the object's displacement. The horizontal axis represents time, t.

The velocity graph shown in Concept 1 is a horizontal line that indicates an object moving at a constant velocity of +5.0 m/s. To determine the object’s displacement between 2.0 and 6.0 seconds from the graph, we calculate the area of the rectangle shown in Concept 1, which equals 20 m.

Why does the area equal the displacement? Because it equals the product of velocity and elapsed time, which is displacement. (Some algebra, using the definition of velocity, yields this result: $v = \bar{v} = \Delta x/\Delta t$, so $\Delta x = v\Delta t$.)

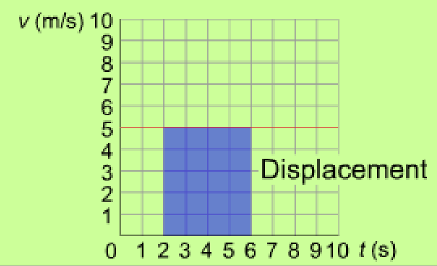
In the second graph, in Concept 2, we show the graph of an object whose velocity changes over time. Again, we can use the principle that the area between the velocity graph and the horizontal time axis equals the displacement. If we know the function describing the velocity, we can calculate the area (and the displacement) exactly using calculus.

Here, we approximate the area by using a set of rectangles. The sum of the areas of the five vertical rectangles between the horizontal axis and the graph is a good approximation for the area under the graph. The area of each rectangle equals the product of a small amount of elapsed time (the width of the rectangle) and a velocity (the height of the graph at that point). By keeping the width of the rectangles small, we ensure that the height of each rectangle is close to the average velocity for that elapsed time.

Since the area is close to the average velocity times the elapsed time, it is a good approximation of the displacement for that period of time. The sum of all the rectangular areas then approximates the total displacement. The sum of the areas of the rectangles in Concept 2 is 20 meters. Using calculus, we calculated the exact displacement as 20.8 meters, so the estimate with rectangles is quite close.

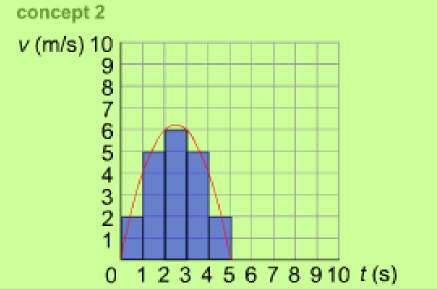
Remember that displacement can be positive or negative. When the velocity graph is above the horizontal axis, the velocity is positive and the displacement is positive. When it is below, the velocity is negative and the displacement during that interval is negative. You see this point emphasized in Concept 3.

concept 1



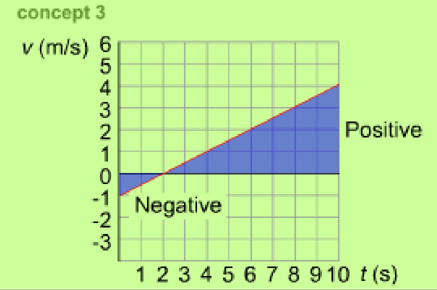
Velocity-time graph
Displacement = area between graph and time axis

concept 2



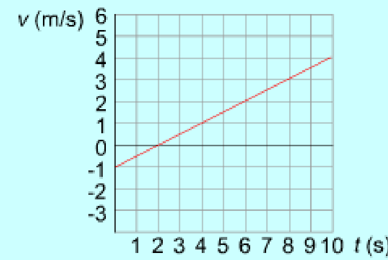
Velocity-time graph
Area approximated by rectangles

concept 3



Velocity-time graph
Negative displacement below time axis

example 1



What is the object's displacement in the first six seconds?

area below t axis = ½(2.0 s)(–1.0 m/s)

area above t axis = ½(4.0 s)(2.0 m/s)

$\Delta x = -1.0\text{ m} + 4.0\text{ m}$

$\Delta x = 3.0\text{ m}$

2.10 - Acceleration

Acceleration: Change in velocity.

A racing car accelerates.

When an object's velocity changes, it accelerates. Acceleration measures the **rate** at which an object speeds up, slows down or changes direction. Any of these variations constitutes a change in velocity. The letter a represents acceleration.

Acceleration is a popular topic in sports car commercials. In the commercials, acceleration is often expressed as how fast a car can go from zero to 60 miles per hour (97 km/h, or 27 m/s). For example, a current model Corvette® automobile can reach 60 mi/h in 4.9 seconds. There are even hotter cars than this in production.

To calculate average acceleration, divide the change in instantaneous velocity by the elapsed time, as shown in Equation 1. To calculate the acceleration of the Corvette, divide its change in velocity, from 0 to 27 m/s, by the elapsed time of 4.9 seconds. The car accelerates at an average rate of 5.5 m/s per second. We typically express this as 5.5 meters per second squared, or 5.5 m/s^2. (This equals 18 ft/s^2, and with this observation we will cease stating values in both measurement systems, in order to simplify the expression of numbers.) Acceleration is measured in units of length divided by time squared. Meters per second squared (m/s^2) express acceleration in SI units.

Let's assume the car accelerates at a constant rate; this means that each second the Corvette moves 5.5 m/s faster. At one second, it is moving at 5.5 m/s; at two seconds, 11 m/s; at three seconds, 16.5 m/s; and so forth. The car's velocity increases by 5.5 m/s every second.

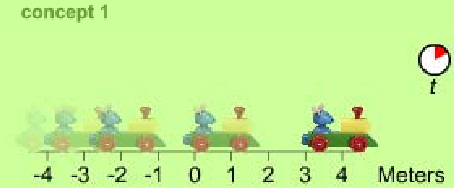
Since acceleration measures the change in **velocity**, an object can accelerate even while it is moving at a constant **speed**. For instance, consider a car moving around a curve. Even if the car's speed remains constant, it accelerates because the change in the car's direction means its velocity (speed plus direction) is changing.

Acceleration can be positive or negative. If the Corvette uses its brakes to go from +60 to 0 mi/h in 4.9 seconds, its velocity is decreasing just as fast as it was increasing before. This is an example of negative acceleration.

You may want to think of negative acceleration as "slowing down," but be careful! Let's say a train has an initial velocity of **negative** 25 m/s and that changes to **negative** 50 m/s. The train is moving at a faster rate (speeding up) but it has negative acceleration. To be precise, its negative acceleration causes an increasingly negative velocity.

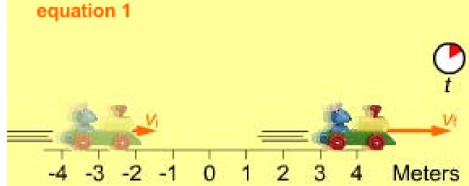
Velocity and acceleration are related but distinct values for an object. For example, an object can have **positive** velocity and **negative** acceleration. In this case, it is slowing down. An object can have zero velocity, yet be accelerating. For example, when a ball bounces off the ground, it experiences a moment of zero velocity as its velocity changes

concept 1



Acceleration
Change in velocity

equation 1



Acceleration

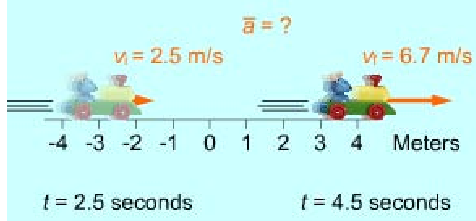
$$\bar{a} = \Delta v / \Delta t$$

$\bar{a}$ = (average) acceleration

from negative to positive, yet it is accelerating at this moment since its velocity is changing.

Δv = change in instantaneous velocity
Δt = elapsed time
Units: meters per second squared (m/s^2)

example 1



What is the average acceleration of the mouse between 2.5 and 4.5 seconds?

$\bar{a} = \Delta v / \Delta t$

$\bar{a} = (6.7 \text{ m/s} - 2.5 \text{ m/s}) / (4.5 \text{ s} - 2.5 \text{ s})$

$\bar{a} = (4.2 \text{ m/s}) / (2.0 \text{ s})$

$\bar{a} = 2.1 \text{ m/s}^2$

2.11 - Average acceleration

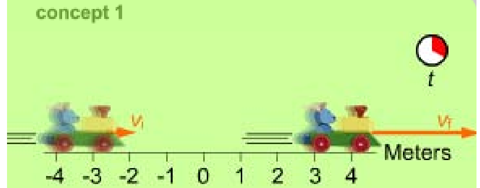
Average acceleration: The change in instantaneous velocity divided by the elapsed time.

Average acceleration is the change in instantaneous velocity over a period of elapsed time. Its definition is shown in Equation 1 to the right.

We will illustrate average acceleration with an example. Let's say you are initially driving a car at 12 meters per second and 8 seconds later you are moving at 16 m/s. The change in velocity is 4 m/s during that time; the elapsed time is eight seconds. Dividing the change in velocity by the elapsed time determines that the car accelerates at an average rate of 0.5 m/s^2.

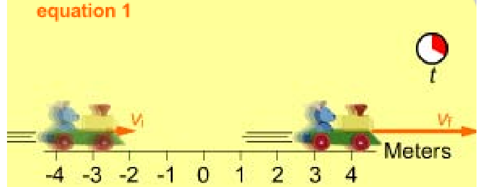
Perhaps the car's acceleration was greater during the first four seconds and less during the last four seconds, or perhaps it was constant the entire eight seconds. Whatever the case, the average acceleration is the same, since it is defined using the initial and final instantaneous velocities.

concept 1



Average acceleration
Change in instantaneous velocity divided by elapsed time

equation 1

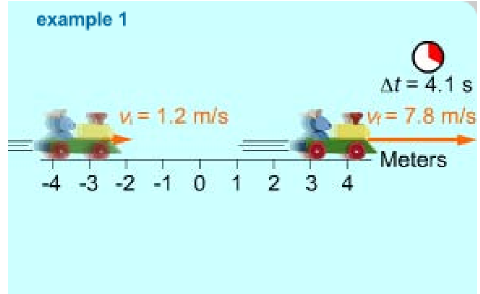


Average acceleration

$$\boxed{\bar{a} = \Delta v / \Delta t}$$

$\bar{a}$ = average acceleration
Δv = change in instantaneous velocity
Δt = elapsed time

example 1



What is the mouse's average acceleration?

$\bar{a} = \Delta v / \Delta t$

$\bar{a} = (7.8 \text{ m/s} - 1.2 \text{ m/s}) / (4.1 \text{ s})$

$\bar{a} = 1.6 \text{ m/s}^2$

2.12 - Instantaneous acceleration

Instantaneous acceleration: Acceleration at a particular moment.

You have learned that velocity can be either average or instantaneous. Similarly, you can determine the average acceleration or the instantaneous acceleration of an object.

We use the mouse in Concept 1 on the right to show the distinction between the two. The mouse moves toward the trap and then wisely turns around to retreat in a hurry. The illustration shows the mouse as it moves toward and then hurries away from the trap. It starts from a rest position and moves to the right with increasingly positive velocity, which means it has a positive acceleration for an interval of time. Then it slows to a stop when it sees the trap, and its positive velocity decreases to zero (this is negative acceleration). It then moves back to the left with increasingly negative velocity (negative acceleration again). If you would like to see this action occur again in the Concept 1 graphic, press the refresh button in your browser.

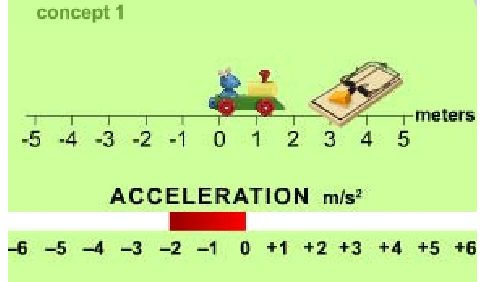
We could calculate an average acceleration, but describing the mouse's motion with instantaneous acceleration is a more informative description of that motion. At some instants in time, it has positive acceleration and at other instants, negative acceleration. By knowing its acceleration and its velocity at an instant in time, we can determine whether it is moving toward the trap with increasingly positive velocity, slowing its rate of approach, or moving away with increasingly negative velocity.

Instantaneous acceleration is defined as the change in velocity divided by the elapsed time as the elapsed time approaches zero. This concept is stated mathematically in Equation 1 on the right.

Earlier, we discussed how the slope of the tangent at any point on a position-time graph equals the instantaneous velocity at that point. We can apply similar reasoning here to conclude that the instantaneous acceleration at any point on a velocity-time graph equals the slope of the tangent, as shown in Equation 2. Why? Because slope equals the rate of change, and acceleration is the rate of change of velocity.

In Example 1, we show a graph of the velocity of the mouse as it approaches the trap and then flees. You are asked to determine the sign of the instantaneous acceleration at four points; you can do so by considering the slope of the tangent to the velocity graph at each point.

concept 1



Instantaneous acceleration
Acceleration at a particular moment

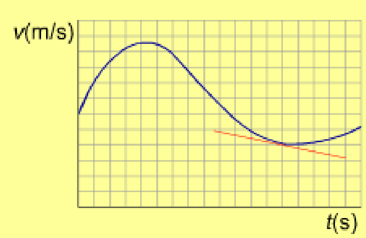
equation 1

Instantaneous acceleration

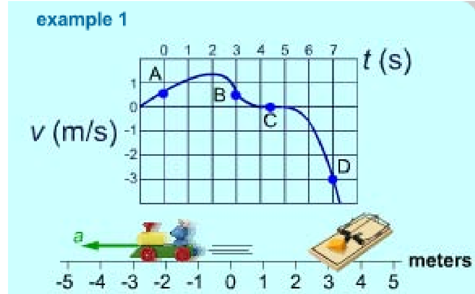
$$a = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t}$$

a = instantaneous acceleration
Δv = change in velocity
Δt = elapsed time (approaches 0)

equation 2



Instantaneous acceleration
Slope of line tangent to point on velocity-time graph



The graph shows the mouse's velocity versus time. Describe the instantaneous acceleration at A, B, C and D as positive, negative or zero.

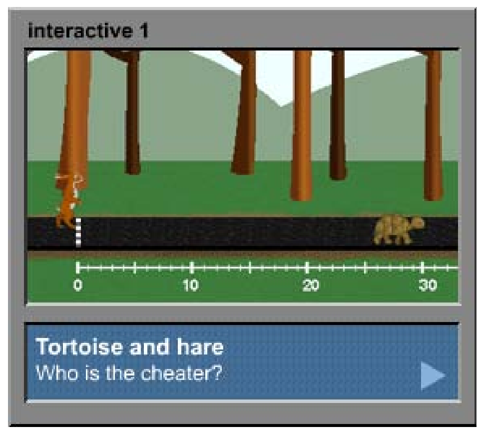
a positive at A

a negative at B

a zero at C

a negative at D

2.13 - Interactive problem: tortoise and hare scandal



The officials at a race have learned that one of the contestants has been eating illegal performance-enhancing vegetables. If either animal exceeds its previous records, it has cheated. Your job is to identify the cheater.

It is well known among track enthusiasts that the tortoise always runs at a constant **velocity** and the hare always runs with a constant **acceleration**. The tortoise's personal best is 5.57 m/s. The hare's most rapid acceleration has been 7.88 m/s^2.

If either animal beats its previous best, it is the guilty party in the vegetable scandal. The hare gives the tortoise a head start of 31.0 meters to make the race more interesting.

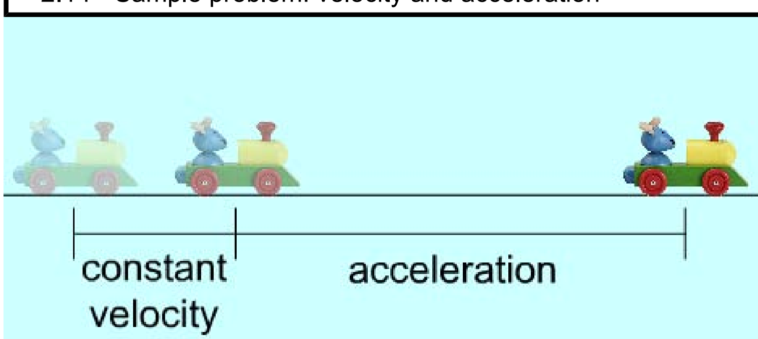
Click on the graphic on the right and press GO to start the race. You will see values for the positions of the tortoise and the velocity of the hare in the control panel. You will need to record these values at a couple times in the race. Press PAUSE during the race to do so. Press PAUSE to resume the race as well.

From the values you record, calculate the velocity of the tortoise and the acceleration of the hare. Who is cheating?

If you decide the tortoise cheated, click on the lettuce; if the culprit is the hare, click on the carrots. The simulation will confirm (or reject) your conclusion. Press RESET and GO to run the race again if you need to.

If you have trouble answering the question, refer to the sections on velocity and acceleration, and look at the equations that define these terms.

2.14 - Sample problem: velocity and acceleration

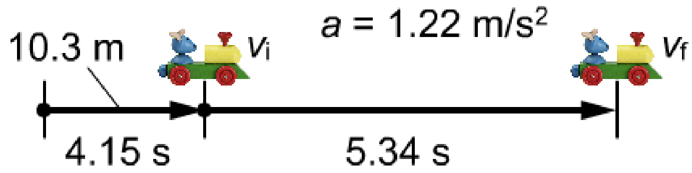


The mouse car goes 10.3 meters in 4.15 seconds at a constant velocity, then accelerates at 1.22 m/s^2 for 5.34 more seconds.

What is its final velocity?

Solving this problem requires two calculations. The mouse car's velocity during the first part of its journey must be calculated. Using that value as the initial velocity of the second part of the journey, and the rate of acceleration during that part, you can calculate the final velocity.

Draw a diagram



Variables

Part 1: Constant velocity

displacement	$\Delta x = 10.3$ m
elapsed time	$\Delta t = 4.15$ s
velocity	v

Part 2: Constant acceleration

initial velocity	$v_i = v$ (calculated above)
acceleration	$a = 1.22$ m/s^2
elapsed time	$\Delta t = 5.34$ s
final velocity	v_f

What is the strategy?

1. Use the definition of velocity to find the velocity of the mouse car before it accelerates. The velocity is constant during the first part of the journey.
2. Use the definition of acceleration and solve for the final velocity.

Physics principles and equations

The definitions of velocity and acceleration will prove useful. The velocity and acceleration are constant in this problem. In this and later problems, we use the definitions for average velocity and acceleration without the bars over the variables.

$$v = \Delta x/\Delta t$$

$$a = \Delta v/\Delta t = (v_f - v_i)/\Delta t$$

Step-by-step solution

We start by finding the velocity before the engine fires.

Step	Reason
1. $v = \Delta x/\Delta t$	definition of velocity
2. $v = (10.3\text{ m})/(4.15\text{ s})$	enter values
3. $v = 2.48$ m/s	divide

Next we find the final velocity using the definition of acceleration. The initial velocity is the same as the velocity we just calculated.

Step	Reason
4. $a = (v_f - v_i)/\Delta t$	definition of acceleration
5. $1.22\text{ m/s}^2 = \dfrac{v_f - 2.48\text{ m/s}}{5.34\text{ s}}$	enter given values, and velocity from step 3
6. $6.51\text{ m/s} = v_f - 2.48\text{ m/s}$	multiply by 5.34 s
7. $v_f = 8.99$ m/s	solve for v_f

2.15 - Interactive checkpoint: subway train

A subway train accelerates along a straight track at a constant 1.90 m/s^2. How long does it take the train to increase its speed from 4.47 m/s to 13.4 m/s?

Answer:

$\Delta t =$ ______ s

2.16 - Interactive problem: what's wrong with the rabbits?

You just bought five rabbits. They were supposed to be constant acceleration rabbits, but you worry that some are the less expensive, non-constant acceleration rabbits. In fact, you think two might be the cheaper critters.

You take them home. When you press GO, they will run or jump for five seconds (well, one just sits still) and then the simulation stops. You can press GO as many times as you like and use the PAUSE button as well.

Your mission: Determine if you were ripped off, and drag the "½ off" sale tags to the cheaper rabbits. The simulation will let you know if you are correct. You may decide to keep the cuddly creatures, but you want to be fairly charged.

Each rabbit has a velocity gauge that you can use to monitor its motion in the simulation. The simplest way to solve this problem is to consider the rabbits one at a time: look at a rabbit's velocity gauge and determine if the velocity is changing at a constant rate. No detailed mathematical calculations are required to solve this problem.

If you find this simulation challenging, focus on the relationship between acceleration and velocity. With a constant rate of acceleration, the velocity must change at a constant rate: no jumps or sudden changes. Hint: No change in velocity is zero acceleration, a constant rate.

2.17 - Derivation: creating new equations

Other sections in this chapter introduced some of the fundamental equations of motion. These equations defined fundamental concepts; for example, average velocity equals the change in position divided by elapsed time.

Several other helpful equations can be derived from these basic equations. These equations enable you to predict an object's motion without knowing all the details. In this section, we derive the formula shown in Equation 1, which is used to calculate an object's final velocity when its initial velocity, acceleration and displacement are known, but **not** the elapsed time. If the elapsed time were known, then the final velocity could be calculated using the definition of velocity, but it is not.

This equation is valid when the acceleration is constant, an assumption that is used in many problems you will be posed.

equation 1

Deriving a motion equation

$$v_f^2 = v_i^2 + 2a\Delta x$$

v_i = initial velocity
v_f = final velocity
a = constant acceleration
Δx = displacement

Variables

We use t instead of Δt to indicate the elapsed time. This is simpler notation, and we will use it often.

acceleration (constant)	a
initial velocity	v_i
final velocity	v_f
elapsed time	t
displacement	Δx

Strategy

First, we will discuss our strategy for this derivation. That is, we will describe our overall plan of attack. These strategy points outline the major steps of the derivation.

1. We start with the definition of acceleration and rearrange it. It includes the terms for initial and final velocity, as well as elapsed time.
2. We derive another equation involving time that can be used to eliminate the time variable from the acceleration equation. The condition of constant acceleration will be crucial here.
3. We eliminate the time variable from the acceleration equation and simplify. This results in an equation that depends on other variables, but not time.

Physics principles and equations

Since the acceleration is constant, the velocity increases at a constant rate. This means the average velocity is the sum of the initial and final velocities divided by two.

$$\bar{v} = (v_i + v_f)/2$$

We will use the definition of acceleration,

$$a = (v_f - v_i)/t$$

We will also use the definition of average velocity,

$$\bar{v} = \Delta x / t$$

Step-by-step derivation

We start the derivation with the definition of average acceleration, solve it for the final velocity and do some algebra. This creates an equation with the square of the final velocity on the left side, where it appears in the equation we want to derive.

Step	Reason
1. $a = (v_f - v_i)/t$	definition of average acceleration
2. $v_f = v_i + at$	solve for final velocity
3. $v_f^2 = (v_i + at)^2$	square both sides
4. $v_f^2 = v_i^2 + 2v_i at + a^2t^2$	expand right side
5. $v_f^2 = v_i^2 + at(2v_i + at)$	factor out at
6. $v_f^2 = v_i^2 + at(v_i + v_i + at)$	rewrite $2v_i$ as a sum
7. $v_f^2 = v_i^2 + at(v_i + v_f)$	substitution from equation 2

The equation we just found is the basic equation from which we will derive the desired motion equation. But it still involves the time variable t – multiplied by a sum of velocities. In the **next** stage of the derivation, we use two different ways of expressing the average velocity to develop a second equation involving time multiplied by velocities. We will subsequently use that second equation to eliminate time from the equation above.

Step	Reason
8. $\bar{v} = \frac{v_i + v_f}{2}$	average velocity is average of initial and final velocities
9. $\bar{v} = \frac{\Delta x}{t}$	definition of average velocity
10. $\frac{v_i + v_f}{2} = \frac{\Delta x}{t}$	set right sides of 8 and 9 equal
11. $t(v_i + v_f) = 2\Delta x$	rearrange equation

We have now developed two equations that involve time multiplied by a sum of velocities. The left side of the equation in step 11 matches an expression appearing in equation 7, at the end of the first stage. By substituting from this equation into equation 7, we eliminate the time variable *t* and derive the desired equation.

Step	Reason
12. $v_f^2 = v_i^2 + a(2\Delta x)$	substitute right side of 11 into 7
13. $v_f^2 = v_i^2 + 2a\Delta x$	rearrange factors

We have now accomplished our goal. We can calculate the final velocity of an object when we know its initial velocity, its acceleration and its displacement, but do not know the elapsed time. The derivation is finished.

2.18 - Motion equations for constant acceleration

$$v_f = v_i + at$$

$$\Delta x = v_i t + \frac{1}{2}at^2$$

$$v_f^2 = v_i^2 + 2a\Delta x$$

$$\Delta x = \frac{1}{2}(v_i + v_f)t$$

Δx = displacement, v = velocity, a = acceleration, t = elapsed time

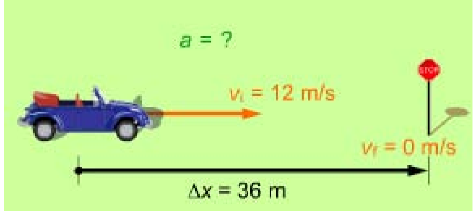
The equations above can be derived from the fundamental definitions of motion (equations such as $a = \Delta v/\Delta t$). To understand the equations, you need to remember the notation: Δx for displacement, v for velocity and a for acceleration. The subscripts i and f represent initial and final values. We follow a common convention here by using t for elapsed time instead of Δt. We show the equations above and below on the right.

Note that to hold true these equations all require a constant rate of acceleration. Analyzing motion with a varying rate of acceleration is a more challenging task. When we refer to acceleration in problems, we mean a constant rate of acceleration unless we explicitly state otherwise.

To solve problems using motion equations like these, you look for an equation that includes the values you know, and the one you are solving for. This means you can solve for the unknown variable.

In the example problem to the right, you are asked to determine the acceleration required to stop a car that is moving at 12 meters per second in a distance of 36 meters. In this problem, you know the initial velocity, the final velocity (stopped = 0.0 m/s) and the displacement. You do not know the elapsed time. The third motion equation includes the two velocities, the acceleration, and the displacement, but does not include the time. Since this equation includes only one value you do not know, it is the appropriate equation to choose.

concept 1



Applying motion equations
Determine the "knowns" and the "unknown(s)"
· Find other knowns from situation
Choose an equation with those variables

equation 1

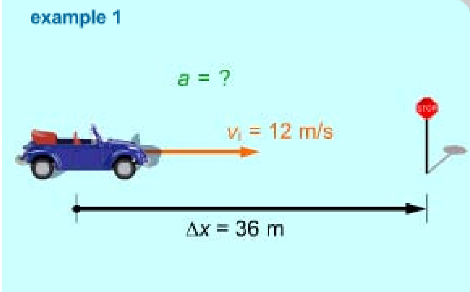
$$v_f = v_i + at$$

$$\Delta x = v_i t + \frac{1}{2}at^2$$

$$v_f^2 = v_i^2 + 2a\Delta x$$

$$\Delta x = \frac{1}{2}(v_i + v_f)t$$

Motion equations



What acceleration will stop the car exactly at the stop sign?

$v_f^2 = v_i^2 + 2a\Delta x$

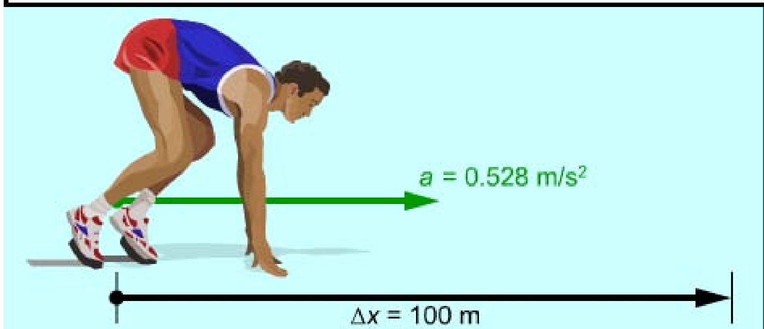
$a = (v_f^2 - v_i^2)/2\Delta x$

$$a = \frac{(0.0 \text{ m/s})^2 - (12 \text{ m/s})^2}{2(36 \text{ m})}$$

$a = -144/72 \text{ m/s}^2$

$a = -2.0 \text{ m/s}^2$

2.19 - Sample problem: a sprinter

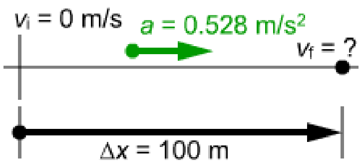


What is the runner's velocity at the end of a 100-meter dash?

You are asked to calculate the final velocity of a sprinter running a 100-meter dash. List the variables that you know and the one you are asked for, and then consider which equation you might use to solve the problem. You want an equation with just one unknown variable, which in this problem is the final velocity.

The sprinter's initial velocity is not explicitly stated, but he starts motionless, so it is zero m/s.

Draw a diagram



Variables

displacement	$\Delta x = 100$ m
acceleration	$a = 0.528 \text{ m/s}^2$
initial velocity	$v_i = 0.00$ m/s
final velocity	v_f

What is the strategy?

1. Choose an appropriate equation based on the values you know and the one you want to find.
2. Enter the known values and solve for the final velocity.

Physics principles and equations

Based on the known and unknown values, the equation below is appropriate. We know all the variables in the equation except the one we are asked to find, so we can solve for it.

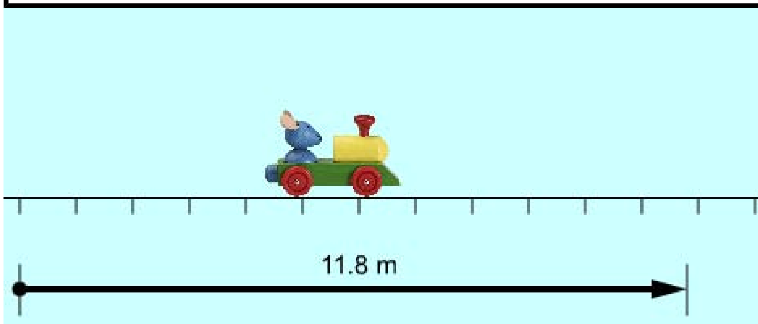
$$v_f^2 = v_i^2 + 2a\Delta x$$

Step-by-step solution

Step	Reason
1. $v_f^2 = v_i^2 + 2a\Delta x$	motion equation
2. $v_f^2 = (0.00\ \text{m/s})^2 + 2(0.528\ \text{m/s}^2)(100\ \text{m})$	enter known values
3. $v_f^2 = 106\ \text{m}^2/\text{s}^2$	multiplication and addition
4. $v_f = 10.3\ \text{m/s}$	take square root

In step 4, we take the square root of 106 to find the final velocity. We chose the positive square root, since the runner is moving in the positive direction. When there are multiple roots, you look at the problem to determine the solution that makes sense given the circumstances. If the runner were running to the left, then a negative velocity would be the appropriate choice.

2.20 - Sample problem: initial and final velocity



The mouse goes 11.8 meters in 3.14 seconds at a constant acceleration of 1.21 m/s^2.

What is its velocity at the beginning and end of the 11.8 meters?

To solve this problem, list the known variables and the ones you are asked for. Since there are two unknown values, the initial and final velocity, you will need to use two equations to solve this problem.

Variables

displacement	$\Delta x = 11.8$ m
acceleration	$a = 1.21$ m/s^2
elapsed time	$t = 3.14$ s
initial velocity	v_i
final velocity	v_f

What is the strategy?

1. There are **two** unknowns, the initial and final velocities, so choose **two** equations that include these two unknowns and the values you do know.
2. Substitute known values and use algebra to reduce the two equations to one equation with a single unknown value.
3. Solve the reduced equation for one of the unknown values, and then calculate the other value.

Physics principles and equations

These two motion equations contain the known and unknown values, and no other values.

$$v_f = v_i + at$$
$$\Delta x = ½(v_i + v_f)t$$

Step-by-step solution

We start with a motion equation containing the two velocities we want to find, and substitute known values, and simplify.

Step	Reason
1. $v_f = v_i + at$	first motion equation
2. $v_f = v_i + (1.21\ \text{m/s}^2)(3.14\ \text{s})$	substitute values
3. $v_f = v_i + 3.80\ \text{m/s}$	multiply

Now we use a second motion equation containing the two velocities, substitute known values, and simplify. This gives us two equations with the two unknowns we want to find.

Step	Reason
4. $\Delta x = \frac{1}{2}(v_i + v_f)t$	second motion equation
5. $11.8\text{ m} = \frac{1}{2}(v_i + v_f)(3.14\text{ s})$	substitute known values
6. $7.52\text{ m/s} = v_i + v_f$	multiply by 2, divide by 3.14 s

Now we solve the two equations.

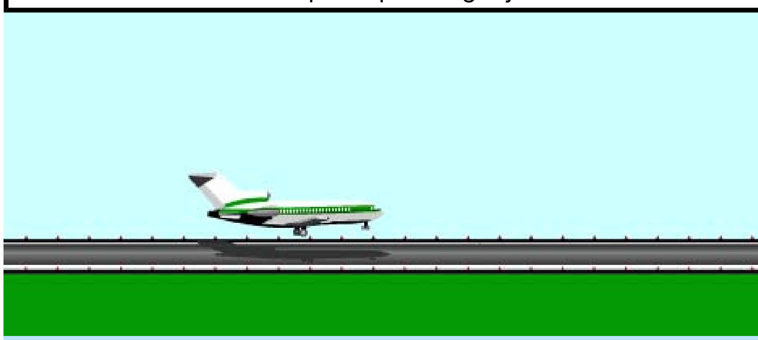
Step	Reason
7. $7.52\text{ m/s} = v_i + v_i + 3.80\text{ m/s}$	substitute equation 3 into equation 6
8. $v_i = 1.86\text{ m/s}$	solve for v_i
9. $v_f = v_i + 3.80\text{ m/s} = 5.66\text{ m/s}$	from equation 3

There are other ways to solve this problem. For example, you could use the equation

$$\Delta x = v_i t + \tfrac{1}{2}at^2$$

to find the initial velocity from the displacement, acceleration, and elapsed time. Then you could use the equation $v_f = v_i + at$ to solve for the final velocity.

2.21 - Interactive checkpoint: passenger jet



A passenger jet lands on a runway with a velocity of 71.5 m/s. Once it touches down, it accelerates at a constant rate of −3.17 m/s^2.
How far does the plane travel down the runway before its velocity is decreased to 2.00 m/s, its taxi speed to the landing gate?

Answer:

$\Delta x =$ ______ m

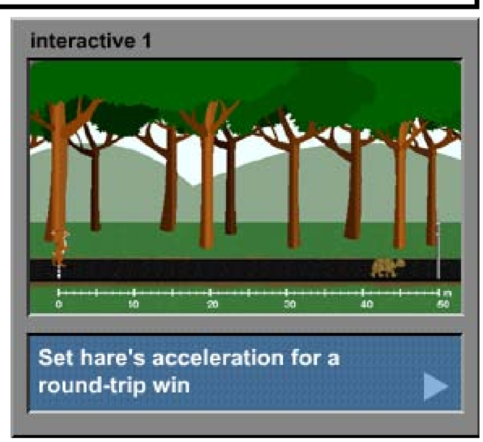
2.22 - Interactive problem: tortoise and hare meet again

The tortoise and the hare are at it again. This time, the race is a round trip. The runners have to go out, turn around a post, and return to the starting point. Your mission is to make the hare win.

The tortoise has a head start of 40.0 meters and is sticking to its strategy of moving at a constant speed. The hare's strategy is to start with a large positive velocity and accelerate so it turns around after passing the post. The hare's initial velocity is 19.97 m/s. The post is 50.0 meters from the starting line, and the hare needs one meter for its turn, so it needs to turn around at 51.0 m. To put it another way: It needs to have zero velocity at this point.

Click on the graphic on the right and set the hare's acceleration to the nearest 0.01 m/s^2, then press GO to see the race. Press RESET to try again.

If you have difficulty getting the hare to win, refer back to the motion equations section.



Free-fall acceleration: Rate of acceleration due to the force of Earth's gravity.

Galileo Galilei is reputed to have conducted an interesting experiment several hundred years ago. According to legend, he dropped two balls with different masses off the Leaning Tower of Pisa and found that both landed at the same time. Their differing masses did not change the time it took them to fall. (We say he was “reputed to have” because there is little evidence that he in fact conducted this experiment. He was more of a “roll balls down a plane” experimenter.)

Today this experiment is used to demonstrate that free-fall acceleration is constant: that the acceleration of a falling object due solely to the force of gravity is constant, regardless of the object’s mass or density. The two balls landed at the same time because they started with the same initial velocity, traveled the same distance and accelerated at the same rate. In 1971, the commander of Apollo 15 conducted a version of the experiment on the Moon, and demonstrated that in the absence of air resistance, a hammer and a feather accelerated at the same rate and reached the surface at the same moment.

In Concept 1, you see a photograph that illustrates free-fall acceleration. Pictures of a freely falling egg were taken every 2/15 of a second. Since the egg’s speed constantly increases, the distance between the images increases over time. Greater displacement over the same interval of time means its velocity is increasing in magnitude; it is accelerating.

Free-fall acceleration is the acceleration caused by the force of the Earth’s gravity, ignoring other factors like air resistance. It is sometimes stated as the rate of acceleration in a vacuum, where there is no air resistance. Near the Earth’s surface, its magnitude is 9.80 meters per second squared. The letter g represents this value. The value of g varies slightly based on location. It is less at the Earth's poles than at the equator, and is also less atop a tall mountain than at sea level.

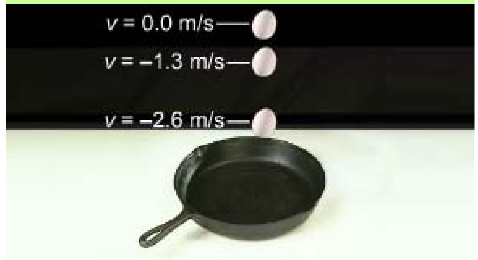
The acceleration of 9.80 m/s^2 occurs in a vacuum. In the Earth’s atmosphere, a feather and a small lead ball dropped from the same height will not land at the same time because the feather, with its greater surface area, experiences more air resistance. Since it has less mass than the ball, gravity exerts less force on it to overcome the larger air resistance. The acceleration will also be different with two objects of the same mass but different surface areas: A flat sheet of paper will take longer to reach the ground than the same sheet crumpled up into a ball.

By convention, “up” is positive, and “down” is negative, like the values on the y axis of a graph. This means when using g in problems, we state free-fall acceleration as **negative** 9.80 m/s^2. To make this distinction, we typically use a or a_y when we are using the negative sign to indicate the direction of free-fall acceleration.

Free-fall acceleration occurs regardless of the direction in which an object is moving. For example, if you throw a ball straight up in the air, it will slow down, accelerating at −9.80 m/s^2 until it reaches zero velocity. At that point, it will then begin to fall back toward the ground and continue to accelerate toward the ground at the same rate. This means its velocity will become increasingly negative as it moves back toward the ground.

The two example problems in this section stress these points. For instance, Example 2 on the right asks you to calculate how long it will take a ball thrown up into the air to reach its zero velocity point (the peak of its motion) and its acceleration at that point.

concept 1



Free-fall acceleration
Acceleration due to gravity

concept 2

Galileo's famous experiment
Confirmed by Apollo 15 on the Moon

equation 1

Free-fall acceleration on Earth

$$g = 9.80 \text{ m/s}^2$$

g = magnitude of free-fall acceleration

example 1

What is the egg's velocity after falling from rest for 0.10 seconds?

$v_f = v_i + at$

$v_f = (0\text{ m/s}) + (–9.80\text{ m/s}^2)(0.10\text{ s})$
$v_f = –0.98\text{ m/s}$

example 2

How long will it take the ball to reach its peak? What is its acceleration at that point?

$v_f = v_i + at$

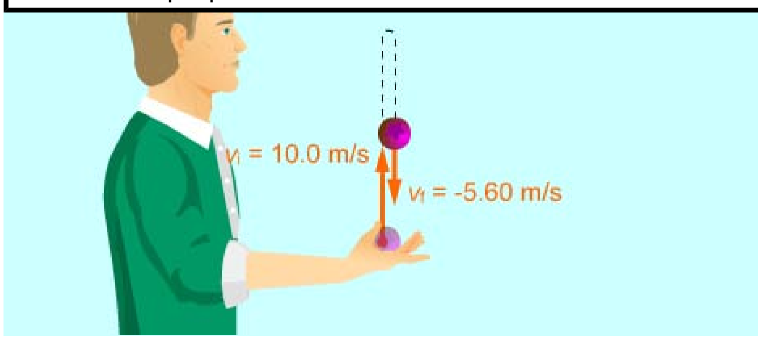
$t = (v_f – v_i)/a$

$t = (0\text{ m/s} – 4.9\text{ m/s})/(–9.80\text{ m/s}^2)$

$t = 4.9/9.80\text{ s} = 0.50\text{ s}$

acceleration = $–9.80\text{ m/s}^2$

2.24 - Sample problem: free fall

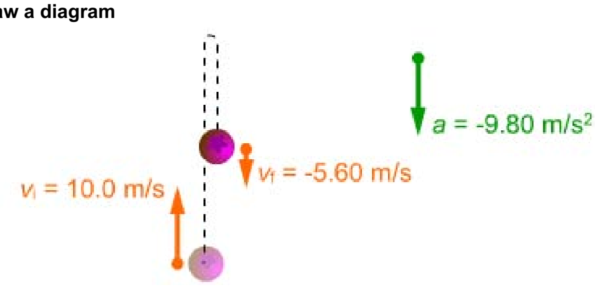


The ball is thrown straight up, with velocity 10.0 m/s.

When will its velocity be −5.60 m/s?

As you see above, a ball is tossed straight up into the air with an initial velocity of positive 10.0 meters per second. You are asked to figure out how long it will take before its velocity is **negative** 5.60 m/s. The ball will have this velocity when it is falling back to the ground.

Draw a diagram



Variables

Be careful with the signs for acceleration and velocity. We use the common convention that upward quantities are positive, and downward negative. The magnitude of the acceleration is the free-fall acceleration constant $g = 9.80\text{ m/s}^2$.

initial velocity	$v_i = 10.0\text{ m/s}$
final velocity	$v_f = −5.60\text{ m/s}$
acceleration	$a = −9.80\text{ m/s}^2$
elapsed time	t

What is the strategy?

1. Choose an appropriate motion equation for the knowns and unknowns.
2. Solve for the elapsed time.

Physics principles and equations

This motion equation involves just the values we know and the time we want to find.

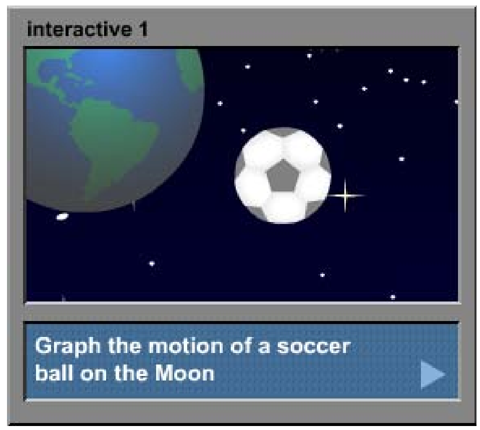
$$v_f = v_i + at$$

Step-by-step solution

A tricky part to solving the problem is making sure the signs for the velocities and acceleration are correct. The acceleration toward the ground is negative, as is the ball's displacement.

Step	Reason
1. $v_f = v_i + at$	motion equation
2. $-5.60 \text{ m/s} = 10.0 \text{ m/s} + (-9.80 \text{ m/s}^2)t$	enter known values
3. $-15.6 \text{ m/s} = (-9.80 \text{ m/s}^2)t$	subtract 10.0
4. $t = 1.60$ s	divide to solve for t

2.25 - Interactive problem: soccer on the Moon

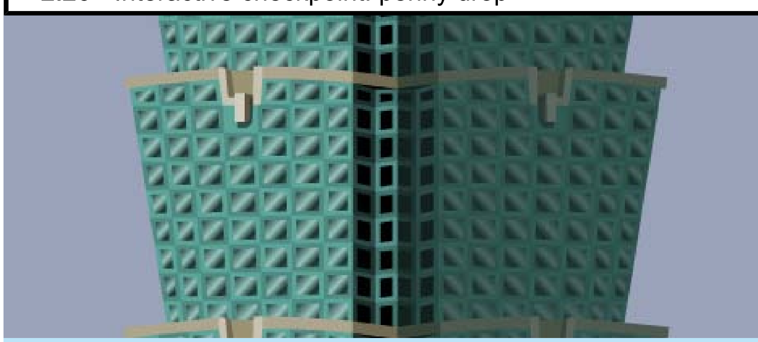


In this simulation, a position-time graph is shown for a soccer ball that is thrown directly upwards and then falls back to the ground. In this case, the action takes place on the Moon, so the free-fall acceleration is not the same as on Earth. Your challenge is to match the graph by setting the soccer ball's initial velocity and the constant free-fall acceleration.

The vertical axis of the graph is the vertical position of the soccer ball, and the horizontal axis is time. The graph goes through exact grid points at the beginning, as the ball reaches its maximum height, and as the ball returns to the ground and stops. These grid point values should help you to calculate both the initial velocity and the acceleration. (You **could** look up the Moon's free-fall acceleration, but there are different reported values, so you may not find the one we used. Better do the math.) Since there are two unknown values, you will need to use two equations. A useful value to know is the velocity of the ball when it reaches its maximum height. You can determine that value without calculation. Review the section on linear motion equations if you are not sure how to start.

Enter the initial velocity to the nearest 0.1 m/s and the acceleration to the nearest 0.1 m/s^2 and press GO. The soccer ball will move, and its position will be graphed. Press RESET to start over.

2.26 - Interactive checkpoint: penny drop



You drop a penny off Taiwan's Taipei 101 tower, which is 509 meters tall. How long does it take to hit the ground? Ignore air resistance, consider down as negative, and the ground as having zero height.

Answer:

$t =$ ______ s

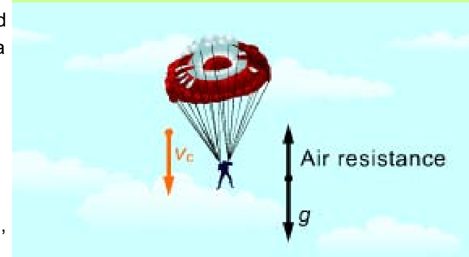
2.27 - Spreadsheet: modeling air resistance

In much of elementary physics, you are told to ignore air resistance. It is complicated to determine; more importantly, it is crucial to first understand motion without air resistance. One of Galileo's great achievements was that he saw through the confusion introduced by air resistance and perceived that the acceleration caused by gravity is constant for all objects. However, air resistance does exist, and it does affect motion.

The parachutist shown in Concept 1 on the right relies on air resistance acting on the parachute to slow his rate of descent. As the parachutist falls, his downward speed increases until the magnitude of the force of air resistance exactly equals that of gravity. At that point, there is no net force acting on him, and he descends at a constant speed, called his *terminal velocity*.

In this section, an equation for air resistance is supplied so you can begin to understand its impact on motion. We use a spreadsheet to calculate the effect of air resistance on a falling object. This provides an introduction to motion with **non-constant** acceleration.

concept 1



Modeling motion with air resistance

Air resistance slows object
· Acceleration proportional to velocity squared
Terminal velocity
· Object reaches constant velocity

Since we are focused on modeling motion, not the theory of air resistance, we will quickly cover a few essential points on how air resistance affects motion. Air resistance is a force, like the force of gravity pulling you down, or the force you exert when you push a door open. In ordinary conditions on Earth, falling objects experience not only the force of gravity, but also this opposing force. The force of air resistance, called drag, acts in a direction opposite to the object's motion. (Note: When an object is rising, air resistance, which always opposes motion, acts in the same direction as gravity, and both act to reduce the object's speed.)

The acceleration due to air resistance "opposes" a falling object's free-fall acceleration and reduces the overall (net) acceleration of the falling object. Unlike free-fall acceleration, the acceleration due to air resistance is not constant. We use a typical equation here for the acceleration due to air resistance. This acceleration is the product of a drag constant k and the **square** of the object's speed. The drag constant depends on the object's shape and surface area. Parachutes are designed to have large drag constants.

equation 1

Modeling motion with air resistance

$$v_c = v_p + (g - kv_p^2)\Delta t$$

Current_velocity = Prior_velocity + (Acceleration – Drag_constant * Prior_velocity^2) * Time_interval

In Equation 1 to the right, you see the same equation in two formats for determining the speed of a falling object. The notation of the first may be a little unfamiliar but it is based on the first of the four standard motion equations shown in a previous section. The second formulation of the equation is the form we use in the spreadsheet that we discuss below.

The equation comes from rearranging the definition of acceleration. It states that the current speed v_c equals the prior speed v_p plus the product of the net acceleration and the elapsed time. (We use the terms "prior" and "current" here instead of "initial" and "final," for reasons we will explain in a moment.)

In this case, the net acceleration is the acceleration caused by gravity **minus** the acceleration in the opposite direction caused by air resistance. The overall equation states that the current speed equals the prior speed plus the change in speed due to the **net** acceleration.

This is not a constant acceleration. The acceleration due to air resistance is a function of the square of the speed, and the speed changes until terminal velocity is reached. The prior speed value is used to calculate the "current" acceleration due to air resistance, which means that for this equation to provide a good approximation, the time increment must be small. This is why we use "current" and "prior" for the speeds: to indicate that they represent the speed in a pair of closely succeeding instants, rather than overall initial and final values.

This is not an easily solved equation (unless you know the right mathematics, namely differential equations). But your computer will let you model the motion without having to solve the differential equation. We used a tool called a spreadsheet to solve the problem. Some well-known spreadsheet programs include Microsoft® Excel, Lotus® 1-2-3™, and the spreadsheets that are part of products like AppleWorks®. You see a portion of the spreadsheet in the illustration on the right.

The spreadsheet solves the problem by dividing the time into small increments. Every 0.01 seconds, the speed and acceleration are recalculated using the formula in Equation 1. Because the time interval is small, this approximates calculating the instantaneous speed and acceleration. It takes 500 iterations to reach 5.0 seconds, but the result is quite accurate.

Depending on your browser, you may be able to launch a spreadsheet program by clicking on the link below. You will then see a spreadsheet calculating the speed at successive instants using the equation described above.

If you open the spreadsheet and click in the middle of the column labeled "Velocity with drag" you will find exactly the second version of the equation shown in Equation 1. We used the spreadsheet's cell naming feature to simplify its expression, instead of the default spreadsheet equations like "=B1−C2."

In the spreadsheet, we set some initial values: the value of g, the initial velocity of the object (0.000 m/s), the increment of time we will use (0.010 s) and the value for k (0.200 m^{-1}). You can vary these. (Note that in the spreadsheet we assign "down" to be the positive direction, for convenience in reading the spreadsheet.)

Let's walk through the first two iterations. After 0.01 seconds, the object's velocity equals its initial velocity (0) plus the product of g and Δt. This equals 0.098 meters per second. The spreadsheet calculates the effect of air resistance as k times the square of prior velocity. Since the prior velocity equals 0, our model says there is no air resistance in the initial 0.01 seconds. This is an approximation in our model since there will be some air resistance, but by choosing a small increment of time we minimize the effect of this approximation. The spreadsheet tells us that after 0.01 seconds, the object moves at 0.098 meters per second.

Now the spreadsheet iterates again. The prior velocity is now 0.098 m/s. This time, air resistance will be a factor. The spreadsheet subtracts the product of k (0.2) and the square of the prior velocity (0.098) from g and multiplies the resulting net acceleration by the increment in time (0.01). The spreadsheet tells us that the velocity after 0.02 seconds is 0.1959808 m/s, a little less than it would be if we had disregarded air resistance. This is displayed as 0.196 by the spreadsheet, so you will not see any difference between the velocity with drag and the velocity without drag yet. But after 0.05 seconds, you can see the velocities with and without air resistance starting to differ.

We could continue, and the spreadsheet does, hundreds of times, thousands if you wish. If you look further down in the spreadsheet, you see that at 3.62 seconds the object reaches a terminal velocity of 7.000 m/s. Not only does this spreadsheet correctly calculate the velocity at any instant in time, it lets you find a value for the terminal velocity when it occurs.

Click here to open the spreadsheet. If the file does not open, on Windows click with the right mouse button and choose the save option. On the Macintosh, hold down the “control” key and click on the link, then choose the option to download the file.

The great thing about spreadsheets is that they do “what if” analysis very well. You may be thinking that 3.62 seconds seems to be a short time to reach terminal velocity, and also wondering what would happen if you used a smaller value for k. You could use the spreadsheet to calculate: if an object reaches a terminal velocity of 50 m/s, how long will it fall before doing so? Just change k until you see that value become the terminal velocity, and observe the elapsed time.

Once you become familiar with spreadsheets, you can build models like this quite quickly. An experienced spreadsheet user could create the model we used here in less than 15 minutes.

Calculating velocity with air resistance

Initial Conditions	
Drag constant	0.200
Time interval	0.010
Acceleration	9.80
Initial velocity	0.000

Calculations

Elapsed time	Velocity w/ drag	Velocity w/o drag
0.01	0.098	0.098
0.02	0.196	0.196
0.03	0.294	0.294
0.04	0.392	0.392
0.05	0.489	0.490
0.06	0.587	0.588
0.07	0.684	0.686
0.08	0.781	0.784
0.09	0.878	0.882
0.10	0.975	0.980
0.11	1.071	1.078
0.12	1.166	1.176
0.13	1.262	1.274
0.14	1.356	1.372
0.15	1.451	1.470
0.16	1.545	1.568
3.58	6.999	35.084
3.59	6.999	35.182
3.60	6.999	35.280
3.61	6.999	35.378
3.62	7.000	35.476
3.63	7.000	35.574
3.64	7.000	35.672
3.65	7.000	35.770
3.66	7.000	35.868
3.67	7.000	35.966
3.68	7.000	36.064
3.69	7.000	36.162
3.70	7.000	36.260
3.71	7.000	36.358

2.28 - Interactive derivation: slamming on the brakes

This chapter has presented many of the commonly used equations for analyzing motion, but more of them can be derived. In this section, we challenge you to derive an equation that is required to solve the problem we pose. You can test its validity in the simulation to the right.

Your spacecraft is collecting the mail from a planet. The mail is fired up from the planet and you pick it up when it has reached its maximum distance from the planet. The mail will be there in 9.00 seconds. Your craft starts 6.00 kilometers from the pickup point. You fire retrorockets to slow down your spacecraft, but alas, your velocity gauge is broken, so you do not know your current velocity. Your spacecraft must have zero velocity when it reaches the mail.

In sum, you know the final desired velocity (zero), the displacement and the time. You set the acceleration to slow the rocket down. If you can develop a motion equation that does not require the initial velocity but includes these other factors, then you can determine what the acceleration should be. Enter the acceleration to the nearest meter per second squared (e.g., 178 m/s^2), then press GO. Press RESET to try again.

For an additional challenge, determine the rate of free-fall acceleration experienced by the mail package. The rate of acceleration is constant, but since the planet in the simulation is not Earth, it does not equal g.

2.29 - Interactive problem: shuffleboard

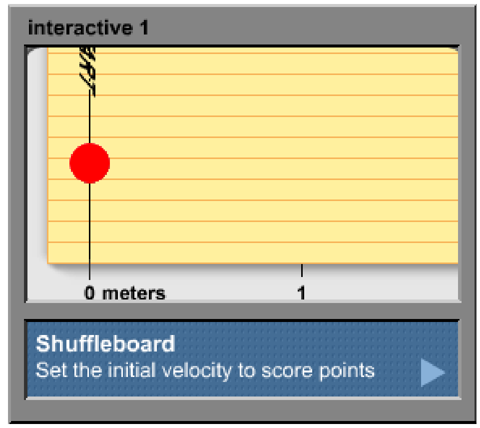
The simulations to the right challenge you to a game that resembles shuffleboard. Your mission is to have the pucks travel exactly 5.00 meters to the goal line on the right edge of the playing board. If you can get a puck to stop on the line in each of the simulations, you will have demonstrated that you have learned the essentials of this chapter and you will win the game.

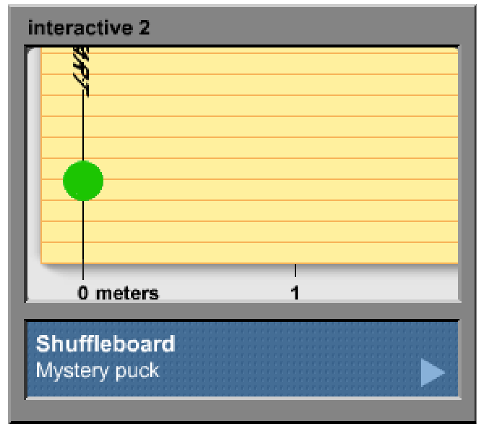
In the first simulation, the red puck has a constant acceleration of −0.289 meters per second squared. Once it reaches zero velocity, it stops moving. (Consider the negative acceleration to be a slowing due to friction.)

Calculate the initial velocity for the puck that will cause it to travel 5.00 meters and stop. Enter this value in the simulation to the nearest 0.01 m/s and press GO to set it moving.

The green puck in the second simulation is a little more mysterious. It also has a constant acceleration, but you do not know what it is. You will have to slide the puck in order to record data and calculate its acceleration. Like the red puck, it will stop when it reaches zero velocity. Pick an initial velocity for the puck and observe what happens − the PAUSE button proves handy as you record data. (An initial velocity of 1.00 m/s is a useful velocity at which to gather data.) Once you calculate the acceleration of the green puck, you can calculate the initial velocity that will cause it to slide 5.00 meters. Enter this value to the nearest 0.01 m/s.

If you have difficulty solving this problem, the section on motion equations will help you.





2.30 - Gotchas

Some errors you might make, or that tests or teachers might try to tempt you to make:

Switching the order in calculating displacement. Remember: It is the final position minus the initial position. If you start at a position of three meters and move to one meter, your displacement is negative two meters. Be sure to subtract three from one, not vice versa.

Confusing distance traveled with displacement. Displacement is the shortest path between the beginning point and final point. It does not matter how the object got there, whether in a straight line or wandering all over through a considerable net distance.

Forgetting the sign. Remember: displacement, velocity and acceleration all have direction. For one-dimensional motion, they require signs indicating the direction. If a problem says that an object moves to the left or down, its displacement is typically negative. Be sure to note the signs of displacement, velocity or acceleration if they are given to you in a problem. Make sure you are consistent with signs. If up is positive, then upward displacement is positive, and the acceleration due to gravity is negative.

Confusing velocity and acceleration. Can an object with zero acceleration have velocity? Yes! A train barreling down the tracks at 150 km/h has velocity. If that velocity is not changing, the train's acceleration is zero.

Can an object with zero velocity have acceleration? Yes again: a ball thrown straight up has zero velocity at the top of its path, but its acceleration at that instant is −9.80 m/s^2.

Confusing constant acceleration with constant velocity. If an object has constant acceleration, it has a constant velocity, right? Quite wrong (unless the constant acceleration is zero). With a constant acceleration other than zero, the velocity is constantly changing.

Misunderstanding negative acceleration. Can something that is "speeding up" also have a negative acceleration? Yes. If something is moving in the negative direction and moving increasingly quickly, it will have a negative velocity and a negative acceleration. When an object has negative velocity and experiences negative acceleration, it will have increasing speed. In other words, negative acceleration is not just "slowing something down." It can also mean an object with negative velocity moving increasingly fast.

2.31 - Summary

Position is the location of an object relative to a reference point called the origin, and is specified by the use of a coordinate system.

Displacement is a measure of the change in the position of an object. It includes both the distance between the object's starting and ending points, and the direction from the starting point to the ending point. An example of displacement would be "three meters west" or "negative two meters".

Similarly, velocity expresses an object's speed **and** direction, as in "three meters per second west." Velocity has a direction. In one dimension, motion in one direction is represented by positive numbers, and motion in the other direction is negative.

An object's velocity may change while it is moving. Its average velocity is its displacement divided by the elapsed time. In contrast, its instantaneous velocity is its velocity at a particular moment. This equals the displacement divided by the elapsed time for a very small interval of time, as the time interval gets smaller and smaller.

Acceleration is a change in velocity. Like velocity, it has a direction and in one dimension, it can be positive or negative. Average acceleration is the change in velocity divided by the elapsed time, and instantaneous acceleration is the acceleration of an object at a specific moment.

Equations

$$\bar{v} = \Delta x / \Delta t$$

$$\bar{a} = \Delta v / \Delta t$$

$$v_f = v_i + at$$

$$\Delta x = v_i t + ½ at^2$$

$$v_f^2 = v_i^2 + 2a\Delta x$$

$$\Delta x = ½ (v_i + v_f)t$$

There are four very useful motion equations for situations where the acceleration is constant. They are the last four equations shown on the right.

Free-fall acceleration, represented by g, is the magnitude of the acceleration due to the force of Earth's gravity. Near the surface of the Earth, falling objects have a downward acceleration due to gravity of 9.80 m/s^2.

Conceptual Problems

C.1 A toddler has become lost in the forest and her father is trying to retrieve her. He is currently located to the north of a large tree and he hears her shouts coming from the south. Do we know from this information whether the toddler is north or south of the tree?

○ Yes ○ No

C.2 Assume Waterville is precisely 100 miles due east of Seattle. (a) With Seattle at the origin, draw a number line that shows only the east-west positions of both cities, with east as positive. (b) Draw a different number line that shows the east-west positions of the cities with Waterville at the origin, and east still positive.

C.3 Julie is a citizen of Country A and she is describing the nearby geography: "Our capital is at the origin. Our famous beaches are 200 km directly west of our capital. Country B's capital is 200 km directly east of our capital. Country B's largest mountain is 700 km to the east of Country A's capital." (a) Draw the four geographical features on a number line from Julie's perspective. (b) A citizen of Country B begs to differ. He believes that his capital is the origin. Draw the features on a number line from his perspective.

C.4 A long straight highway has markers along the side that represent the distance north (in miles) from the start of the highway. In terms of displacement, is there any difference in moving (a) from the "3" marker to the "7" marker, or from "4" to "8"? (b) From "1" to "3", or from "3" to "1"?

(a) ○ Yes ○ No
(b) ○ Yes ○ No

C.5 A diver is standing on a diving board that is 15 meters above the surface of the water. He jumps 0.50 meters straight up into the air, dives into the water and goes 3.0 meters underwater before returning to the surface. (a) Assume that "up" is the positive direction. What is the vertical displacement of the diver from when he is standing on the diving board to when he emerges from underwater? (b) Now assume that "down" is the positive direction. Determine the displacement, as before.

(a) __________ m
(b) __________ m

C.6 Can a car with negative velocity move faster than a car with positive velocity? Explain.

○ Yes ○ No

C.7 Two people drive their cars from Piscataway, New Jersey, to Perkasie, Pennsylvania, about 35 miles east, leaving at the same time and using the same route. The first driver travels at a constant speed the whole trip. The second driver stops for a while for doughnuts and coffee in Lambertville. They arrive in Perkasie at the same time. Do the cars have the same average velocity? Explain.

○ Yes ○ No

C.8 Elaine wants to return a video she rented at the video store, which is 5.0 kilometers away in the positive direction. It takes her 10 minutes to drive to the store, 1.0 minute to deposit the videotape, and 9.0 more minutes to drive home. What is her average velocity for the entire trip?

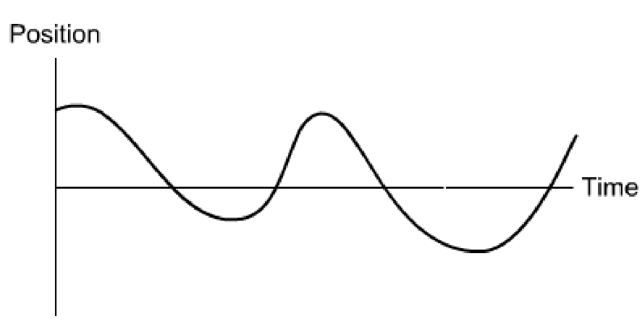
__________ m/s

C.9 An airplane starts out at the Tokyo Narita Airport (x = 0) destined for Bangkok, Thailand (x = 4603 km). After departing from the gate, the plane has to wait ten minutes on the runway before taking off. When the plane is halfway to Bangkok, and cruising at its top speed, which of these is greater: instantaneous velocity, or average velocity since leaving the gate?

i. Instantaneous velocity
ii. Average velocity
iii. They're the same

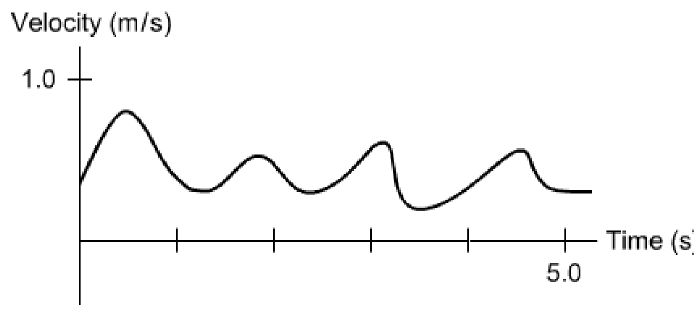
C.10 The position versus time graph for a man trapped on an island is shown. Is he traveling at a constant velocity? Explain.

○ Yes ○ No



C.11 The velocity of a butterfly is shown. For the entire time interval, is the displacement of the butterfly positive, negative, or zero? Explain.

i. Positive
ii. Negative
iii. Zero
iv. Cannot be determined

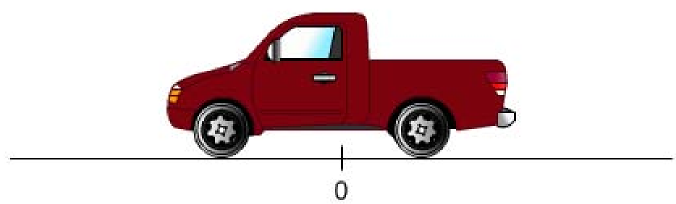


C.12 A penguin swimming through the icy Antarctic waters accelerates from 2.0 m/s to 4.0 m/s over a period of 4.0 seconds, then slows down to a stop over a period of 3.0 seconds and reverses direction, accelerating to 3.0 m/s in the opposite direction over a period of 4.0 seconds. Assume the penguin's acceleration was constant over each of the three time periods, and that it started out traveling in the positive direction. Draw the penguin's velocity versus time graph.

C.13 An eagle is soaring in a straight line from t = 0 s to t = 10 s. Over that time interval, her average velocity is always greater than or equal to her instantaneous velocity. If her initial velocity is 5.0 m/s, draw two different possible velocity versus time graphs for the eagle.

C.14 A truck full of corn is parked at x = 0 and is pointed in the negative direction. If the driver puts it into reverse and holds down on the accelerator, will the following quantities be positive, zero, or negative after one second? (a) position, (b) velocity, (c) acceleration.

(a) i. Positive
ii. Negative
iii. Zero
(b) i. Positive
ii. Negative
iii. Zero
(c) i. Positive
ii. Negative
iii. Zero



C.15 David is driving a minivan to work, and he is stopped at a red light. The light turns green and David drives to the next red light, where he stops again. Is David's average acceleration from light to light positive, negative, or zero?

i. Positive
ii. Negative
iii. Zero

C.17 If you know only the initial position, the final position, and the constant acceleration of an object, can you calculate the final velocity? Explain.

○ Yes ○ No

C.18 A football is thrown vertically up in the air at 10 m/s. If it is later caught at the same spot it was thrown from, will the speed be greater than, less than, or the same as when it was thrown? Ignore air resistance.

i. Greater than
ii. Less than
iii. Same

C.19 Can an object be increasing in speed as the magnitude of its acceleration decreases? If not, explain why not, and if so, provide an example.

○ Yes ○ No

C.20 Acceleration is the change in velocity with respect to time. The change in acceleration with respect to time is sometimes called the "jerk." What are its units?

○ m/s^3
○ m/s^2
○ s/m
○ s/m^2

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following question. What kind of acceleration will cause the hare to change directions?

i. Positive
ii. Negative
iii. Positive or negative

Section 1 - Position

1.1 Anita and Nick are playing tug-of-war near a mud puddle. They are each holding on to an end of a taut rope that has a knot exactly in the middle. Anita's position is 6.2 meters east of the center of the puddle and Nick's position is 3.0 meters west of the center of the puddle. What is the location of the knot relative to the center of the puddle? Treat east as positive and west as negative.

__________ meters

Section 2 - Displacement

2.1 A photographer wants to take a picture of a particularly interesting flower, but he is not sure how far away to place his camera. He takes three steps forward, four back, seven forward, then five back. Finally, he takes the photo. As measured in steps, what was his displacement? Assume the forward direction is positive.

__________ step(s)

2.2 The school bus picks up Brian in front of his house and takes him on a straight-line 2.1 km bus ride to school in the positive direction. He walks home after school. If the front of Brian's house is the origin, (a) what is the position of the school, (b) what is his displacement on the walk home, and (c) what is his displacement due to the combination of the bus journey and his walk home?

(a) __________ km
(b) __________ km
(c) __________ km

2.3 A strange number line is measured in meters to the left of the origin, and in kilometers to the right of the origin. An object moves from −2200 m to 3.1 km. Find its displacement (a) in kilometers, and (b) in meters.

(a) __________ km
(b) __________ m

2.4 The Psychic Squishy Sounds band is on tour in Texas and now are resting in Houston. They traveled 60.0 miles due south from their first concert to Huntsville. If Huntsville is 80.0 miles due north of Houston, what is their displacement from the first concert to Houston? Use the convention that north is positive, and south is negative.

__________ miles

Section 3 - Velocity

3.1 To estimate the distance you are from a lightning strike, you can count the number of seconds between seeing the flash and hearing the associated thunderclap. For this purpose, you can consider the speed of light to be infinite (it arrives instantly). Sound travels at about 343 m/s in air at typical surface conditions. How many kilometers away is a lightning strike for every second you count between the flash and the thunder?

__________ km

3.2 A jogger is moving at a constant velocity of +3.0 m/s directly towards a traffic light that is 100 meters away. If the traffic light is at the origin, x = 0 m, what is her position after running 20 seconds?

__________ m

3.3 A slug has just started to move straight across a busy street in Littletown that is 8.0 meters wide, at a constant speed of 3.3 millimeters per second. The concerned drivers on the street halt until the slug has reached the opposite side. How many seconds elapse until the traffic can start moving again?

__________ s

3.4 Light travels at a constant speed of 3.0×10^8 m/s in a vacuum. (a) It takes light about 1.3 seconds to travel from the Earth to the Moon. Estimate the distance of the Moon from the Earth's surface, in meters. (b) The astronomical unit (abbreviated AU) is equal to the distance between the Earth and the Sun. One AU is about 1.5×10^{11} m. If the Sun suddenly ceased to emit light, how many minutes would elapse until the Earth went dark?

(a) __________ m
(b) __________ min

Section 4 - Average velocity

4.1 In 1271, Marco Polo departed Venice and traveled to Kublai Khan's court near Beijing, approximately 7900 km away in a direction we will call positive. Assume that the Earth is flat (as some did at the time) and that the trip took him 4.0 years, with 365 days in a year. (a) What was his average velocity for the trip, in meters per second? (b) A 767 could make the same trip in about 9.0 hours. What is the average velocity of the plane in meters per second?

(a) __________ m/s
(b) __________ m/s

4.2 An airport shuttle driver is assigned to drive back and forth between a parking lot (located at 0.0 km on a number line) and the airport main terminal (located at +4.0 km). The driver starts out at the terminal at noon, arrives at the lot at 12:15 P.M., returns to the terminal at 12:30 P.M., and arrives back at the lot at 12:45 P.M. What is the average velocity of the shuttle between (a) noon and 12:15, (b) noon and 12:30, and (c) noon and 12:45? Express all answers in meters per second.

(a) __________ m/s
(b) __________ m/s
(c) __________ m/s

4.3 You are driving in one direction on a long straight road. You drive in the positive direction at 126 km/h for 30.0 minutes, at which time you see a police car with someone pulled over, presumably for speeding. You then drive in the same direction at 100 km/h for 45.0 minutes. (a) How far did you drive? (b) What was your average velocity in kilometers per hour?

(a) __________ km
(b) __________ km/h

4.4 You made a journey, and your displacement was +95.0 km. Your initial velocity was +167 km/h and your final velocity was −26.0 km/h. The journey took 43.0 minutes. What was your average velocity in kilometers per hour?

__________ km/h

4.5 The first controlled, sustained flight in a heavier-than-air craft was made by Orville Wright on December 17, 1903. The plane took off at the end of a rail that was 60 feet long, and landed 12 seconds later, 180 feet away from the beginning of the rail. Assume the rail was essentially at the same height as the ground. (a) Calculate the average velocity of the plane in feet per second while it was in the air. (b) What is the average velocity in kilometers per hour?

(a) __________ ft/s
(b) __________ km/h

4.6 A horse is capable of moving at four different speeds: walk (1.9 m/s), trot (5.0 m/s), canter (7.0 m/s), and gallop (12 m/s). Ann is learning how to ride a horse. She spends 15 minutes riding at a walk and 2.4 minutes at each other speed. If she traveled the whole way in the positive direction, what was her average velocity over the trip?

__________ m/s

4.7 A vehicle is speeding at 115 km/h on a straight highway when a police car moving at 145 km/h enters the highway from an onramp and starts chasing it. The speeder is 175 m ahead of the police car when the chase starts, and both cars maintain their speeds. How much time, in seconds, elapses until the police car overtakes the speeder?

__________ s

Section 5 - Instantaneous velocity

5.1 The tortoise and the hare start a race from the same starting line, at the same time. The tortoise moves at a constant 0.200 m/s, and the hare at 5.00 m/s. (a) How far ahead is the hare after five minutes? (b) How long can the hare then snooze until the tortoise catches up?

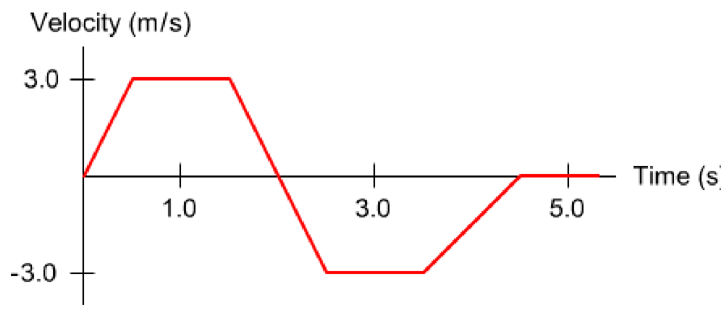
(a) __________ m
(b) __________ s

5.2 You own a yacht which is 14.5 meters long. It is motoring down a canal at 10.6 m/s. Its bow (the front of the boat) is just about to begin passing underneath a bridge that is 30.0 m across. How much time is required until its stern (the end of the boat) is no longer under the bridge?

__________ s

5.3 The velocity versus time graph of a unicycle is shown. What is the instantaneous velocity of the unicycle at (a) t = 1.0 s, (b) t = 3.0 s, and (c) t = 5.0 s?

(a) __________ m/s
(b) __________ m/s
(c) __________ m/s



5.4 Two boats are initially separated by distance d and head directly toward one another. The skippers of the boats want to arrive at the same time at the point that is halfway between their starting points. Boat 1 moves at a speed v and boat 2 moves at twice the speed of boat 1. Because it moves faster, boat 2 starts at time t later than boat 1. The skippers want to know how much later boat 2 should start than boat 1. Provide them with an equation for t in terms of d and v.

- ○ $t = d/4v$
- ○ $t = 3d/2v$
- ○ $t = 3d/4v$
- ○ $t = d/2v$

5.5 This problem requires you to apply some trigonometry. A friend of yours is 51.0 m directly to your left. You and she start running at the same time, and both run in straight lines at constant speeds. You run directly forward at 5.00 m/s for 125 m. She runs to the same final point as you, and wants to arrive at the same moment you do. How fast must she run?

__________ m/s

Section 6 - Position-time graph and velocity

6.1 A fish swims north at 0.25 m/s for 3.0 seconds, stops for 2.0 seconds, and then swims south at 0.50 m/s for 4.0 seconds. Draw a position-time graph of the fish's motion, using north as the positive direction.

6.2 A renegade watermelon starts at x = 0 m and rolls at a constant velocity to x = −3.0 m at time t = 5.0 seconds. Then, it bumps into a wall and stops. Draw its position versus time graph from t = 0 s to t = 10 s.

Section 7 - Interactive problem: draw a position-time graph

7.1 Use the information given in the interactive problem in this section to answer the following questions. Assume that the positive x direction is to the right. (a) Is the ball moving to the left or right from 0 s to 3.0 s? (b) Is the velocity from 3.0 s to 6.0 s positive, negative or zero? (c) What is a good strategy to match the graph from 7.0 s to 10.0 s? Test your answers using the simulation.

(a) i. Left
ii. Right
iii. It is not moving

(b) i. Positive
ii. Negative
iii. Zero

(c) i. Start the ball moving slowly to the left and then increase its speed.
ii. Start the ball quickly moving to the left remaining at a constant speed.
iii. Start the ball slowly moving to the right and then increase its speed.
iv. Start the ball moving quickly to the right and then decrease its speed.

Section 8 - Interactive problem: match a graph using velocity

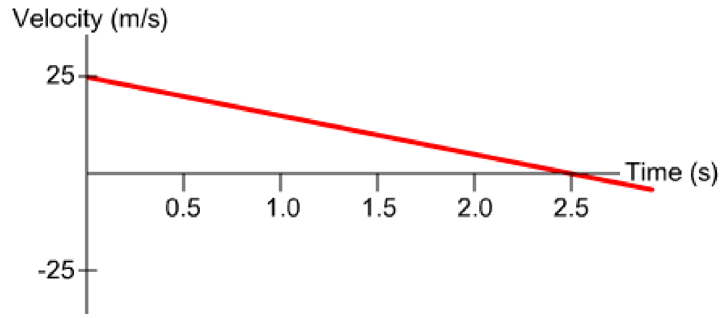
8.1 Use the information given in the interactive problem in this section to answer the following questions. What is the velocity of the ball from (a) 0 to 3.0 seconds, (b) 3.0 to 7.0 seconds, and (c) 7.0 to 10.0 seconds needed to match the graph? Test your answer using the simulation.

(a) __________ m/s
(b) __________ m/s
(c) __________ m/s

Section 9 - Velocity graph and displacement

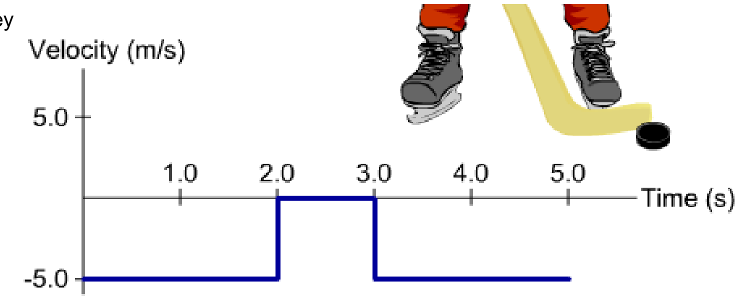
9.1 A clown is shot straight up out of a cannon. The graph of his velocity versus time is shown. Determine the clown's vertical displacement from the instant he is shot out of the cannon at 0.0 seconds until when he reaches zero velocity at 2.5 seconds.

__________ m



9.2 A graph of the velocity versus time of a hockey puck is shown. Calculate the puck's displacement from t = 1.0 s to t = 4.0 s.

__________ m



Section 10 - Acceleration

10.1 The speed limit on a particular freeway is 28.0 m/s (about 101 km/hour). A car that is merging onto the freeway is capable of accelerating at 2.25 m/s^2. If the car is currently traveling forward at 13.0 m/s, what is the shortest amount of time it could take the vehicle to reach the speed limit?

__________ s

10.2 A sailboat is moving across the water at 3.0 m/s. A gust of wind fills its sails and it accelerates at a constant 2.0 m/s^2. At the same instant, a motorboat at rest starts its engines and accelerates at 4.0 m/s^2. After 3.0 seconds have elapsed, find the velocity of (a) the sailboat, and (b) the motorboat.

(a) __________ m/s
(b) __________ m/s

Section 11 - Average acceleration

11.1 A rail gun uses electromagnetic energy to accelerate objects quickly over a short distance. In an experiment, a 2.00 kg projectile remains on the rails of the gun for only 2.10e−2 s, but in that time it goes from rest to a velocity of 4.00×10^3 m/s. What is the average acceleration of the projectile?

__________ m/s^2

11.2 A baseball is moving at a speed of 40.0 m/s toward a baseball player, who swings his bat at it. The ball stays in contact with the bat for 5.00×10^{-4} seconds, then moves in essentially the opposite direction at a speed of 45.0 m/s. What is the magnitude of the ball's average acceleration over the time of contact? (These figures are good estimates for a professional baseball pitcher and batter.)

__________ m/s^2

11.3 A space shuttle sits on the launch pad for 2.0 minutes, and then goes from rest to 4600 m/s in 8.0 minutes. Treat its motion as straight-line motion. What is the average acceleration of the shuttle (a) during the first 2.0 minutes, (b) during the 8.0 minutes the shuttle moves, and (c) during the entire 10 minute period?

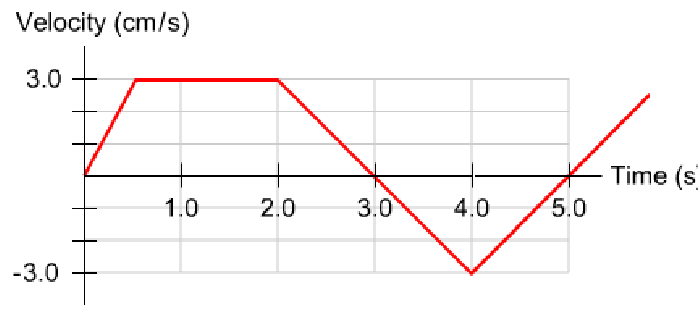
(a) __________ m/s^2
(b) __________ m/s^2
(c) __________ m/s^2

11.4 A particle's initial velocity is −24.0 m/s. Its final velocity, 3.12 seconds later, is −14.0 m/s. What was its average acceleration?

__________ m/s^2

11.5 The velocity versus time graph of an ant is shown. What is the ant's acceleration at (a) t = 1.0 s, (b) t = 3.0 s, and (c) t = 5.0 s?

(a) ________ cm/s^2
(b) ________ cm/s^2
(c) ________ cm/s^2



11.6 The driver of a jet-propelled car is hoping to break the world record for ground-based travel of 341 m/s. The car goes from rest to its top speed in 26.5 seconds, undergoing an average acceleration of 13.0 m/s^2. Upon reaching the top speed, the driver immediately puts on the brakes, causing an average acceleration of −5.00 m/s^2, until the car comes to a stop. (a) What is the car's top speed? (b) How long will the car be in motion?

(a) ________ m/s
(b) ________ s

11.7 A particle's initial velocity is v. Every second, its velocity doubles. Which of the following expressions would you use to calculate its average acceleration after five seconds of motion?

○ $32v/5$ ○ $2v$ ○ $31v$ ○ $31v/5$

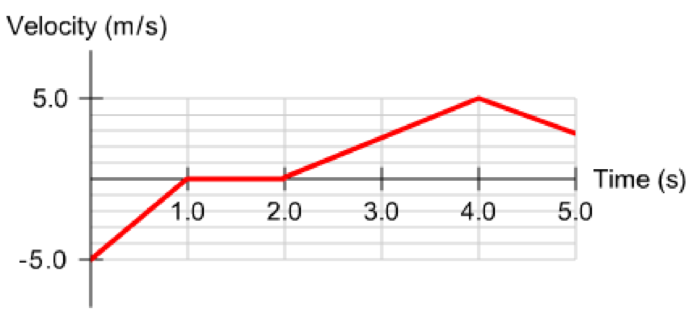
Section 12 - Instantaneous acceleration

12.1 The velocity versus time graph for a pizza delivery driver who is frantically trying to deliver a pizza is shown. (a) During what time interval is he traveling at a constant velocity? (b) During what time interval is his acceleration 5.0 m/s^2? (c) During what time is his acceleration negative?

(a) i. 0 to 1.0 s
ii. 1.0 s to 2.0 s
iii. 2.0 to 4.0 s
iv. During no time interval

(b) i. 0 to 1.0 s
ii. 2.0 to 4.0 s
iii. 4.0 to 5.0 s
iv. During no time interval

(c) i. 0 to 1.0 s
ii. 2.0 to 4.0 s
iii. 4.0 to 5.0 s
iv. During no time interval



Section 13 - Interactive problem: tortoise and hare scandal

13.1 Use the information given in the interactive problem in this section to calculate the tortoise's velocity and the hare's acceleration. Who cheated? Test your answer using the simulation.

○ The tortoise ○ The hare

Section 16 - Interactive problem: what's wrong with the rabbits?

16.1 Use the information given in the interactive problem in this section to find the rabbits that do not have a constant acceleration. Which are the faulty rabbits? Test your answer using the simulation.

☐ The brown rabbit
☐ The tan rabbit
☐ The white rabbit
☐ The grey rabbit
☐ The black rabbit

Section 18 - Motion equations for constant acceleration

18.1 The United States and South Korean soccer teams are playing in the first round of the World Cup. An American kicks the ball, giving it an initial velocity of 3.6 m/s. The ball rolls a distance of 5.0 m and is then intercepted by a South Korean player. If the ball accelerates at −0.50 m/s^2 while rolling along the grass, find its velocity at the time of interception.

__________ m/s

18.2 Which one of the following equations could be used to calculate the time of a journey when the initial velocity is zero, and the constant acceleration and displacement are known?

A. $t = ½a\Delta x$

B. $t = \sqrt{2a\Delta x}$

C. $t = \dfrac{2\Delta x}{a}$

D. $t = \sqrt{\dfrac{2\Delta x}{a}}$

○ A ○ B ○ C ○ D

18.3 The city is trying to figure out how long the traffic light should stay yellow at an intersection. The speed limit on the road is 45.0 km/h and the intersection is 23.0 m wide. A car is traveling at the speed limit in the positive direction and can brake with an acceleration of −5.20 m/s^2. (a) If the car is to stop on the white line, before entering the intersection, what is the minimum distance from the line at which the driver must apply the brakes? (b) How long should the traffic light stay yellow so that if the car is just closer than that minimum distance when the light turns yellow, it can safely cross the intersection without having to speed up?

(a) __________ m
(b) __________ s

18.4 The brochure advertising a sports car states that the car can be moving at 100.0 km/h, and stop in 37.19 meters. What is its average acceleration during a stop from that velocity? Express your answer in m/s^2. Consider the car's initial velocity to be a positive quantity.

__________ m/s^2

18.5 Engineers are designing a rescue vehicle to catch a runaway train. When the rescue vehicle is launched from a stationary position, the train will be at a distance *d* meters away, moving at constant velocity *v* meters per second. The rescue vehicle needs to reach the train in *t* seconds. Write an equation for the constant acceleration needed for the rescue vehicle in terms of *d*, *v*, and *t*.

○ $a = 2(vt + d)/t^2$
○ $a = (vt - d)/t$
○ $a = 2(dv + t)/vt^2$
○ $a = t(v - d)$

18.6 Two spacecraft are 13,500 m apart and moving directly toward each other. The first spacecraft has velocity 525 m/s and accelerates at a constant −15.5 m/s^2. They want to dock, which means they have to arrive at the same position at the same time with zero velocity. (a) What should the initial velocity of the second spacecraft be? (b) What should be its constant acceleration?

(a) __________ m/s
(b) __________ m/s^2

18.7 A Honda® and a Porsche® race, starting from the same point. The Honda accelerates at a constant 4.00 m/s^2; the Porsche at a constant 8.00 m/s^2. The Porsche gives the Honda an advantage by letting it start first. The Honda accelerates, and when it is traveling at 23.0 m/s, the Porsche starts. How far do the cars travel from the starting point before the Porsche catches up with the Honda?

__________ m

18.8 Superguy has attached a nitro booster to his motorcycle to assist him in his quest to rid Science City of his archenemy, Mad Maxwell. The motorcycle doubles its acceleration when the nitro is activated. Superguy is idling by the side of the highway when Maxwell tears by. At this instant Superguy steps on the gas and accelerates at a constant rate until he is moving at the same speed as Maxwell, then uses his booster to close the distance and catch Maxwell. Maxwell moves at a constant 30.0 m/s the entire time, and is caught 60.0 seconds after passing Superguy. (a) How far does Maxwell travel after passing Superguy before being caught? (b) What is the acceleration of the motorcycle with the nitro booster engaged? (c) How long is Superguy in motion before engaging the booster?

(a) __________ m
(b) __________ m/s^2
(c) __________ s

Section 22 - Interactive problem: tortoise and hare meet again

22.1 Use the information given in the interactive problem in this section to calculate the acceleration the hare needs to win the race. Test your answer using the simulation.

__________ m/s^2

Section 23 - Free-fall acceleration

23.1 An elevator manufacturing company is stress-testing a new elevator in an airless test shaft. The elevator is traveling at an unknown velocity when the cable snaps. The elevator falls 1.10 meters before hitting the bottom of the shaft. The elevator was in free fall for 0.900 seconds. Determine its velocity when the cable snapped. As usual, up is the positive direction.

__________ m/s

23.2 A watermelon cannon fires a watermelon vertically up into the air at a velocity of + 9.50 m/s, starting from an initial position 1.20 meters above the ground. When the watermelon reaches the peak of its flight, what is (a) its velocity, (b) its acceleration, (c) the elapsed time, and (d) its height above the ground?

(a) __________ m/s
(b) __________ m/s^2
(c) __________ s
(d) __________ m

23.3 A croissant is dropped from the top of the Eiffel Tower. The height of the tower is 300.5 meters (ignoring the antenna, and this figure changes slightly with temperature). Ignoring air resistance, at what speed will the croissant be traveling when it hits the ground?

__________ m/s

23.4 On a planet that has no atmosphere, a rocket 14.2 m tall is resting on its launch pad. Freefall acceleration on the planet is 4.45 m/s^2. A ball is dropped from the top of the rocket with zero initial velocity. (a) How long does it take to reach the launch pad? (b) What is the speed of the ball just before it reaches the ground?

(a) __________ s
(b) __________ m/s

23.5 To determine freefall acceleration on a moon with no atmosphere, you drop your handkerchief off the roof of a baseball stadium there. The roof is 113 meters tall. The handkerchief reaches the ground in 18.2 seconds. What is freefall acceleration on this moon? (State the result as a positive quantity.)

__________ m/s^2

23.6 You are a bungee jumping fanatic and want to be the first bungee jumper on Jupiter. The length of your bungee cord is 45.0 m. Freefall acceleration on Jupiter is 23.1 m/s^2. What is the ratio of your speed on Jupiter to your speed on Earth when you have dropped 45.0 m? Ignore the effects of air resistance and assume that you start at rest.

____

23.7 On the Apollo 15 space mission, Commander David R. Scott verified Galileo's assertion that objects of different masses accelerate at the same rate. He did so on the Moon, where the acceleration due to gravity is 1.62 m/s^2 and there is no air resistance, by dropping a hammer and a feather at the same time. Assume they were 1.25 meters above the surface of the Moon when he released them. How long did they take to land?

__________ s

23.8 You stand near the edge of Half Dome in Yosemite, reach your arm over the railing, and (thoughtlessly, since what goes up does come down and there are people below) throw a rock upward at 8.00 m/s. Half Dome is 1460 meters high. How long does it take for the rock to reach the ground? Ignore air resistance.

__________ s

23.9 To get a check to bounce 0.010 cm in a vacuum, it must reach the ground moving at a speed of 6.7 m/s. At what velocity toward the ground must you throw it from a height of 1.4 meters in order for it to it have a speed of 6.7 m/s when it reaches the ground? Treat downward as the negative direction, and watch the sign of your answer.

__________ m/s

23.10 Two rocks are thrown off the edge of a cliff that is 15.0 m above the ground. The first rock is thrown upward, at a velocity of +12.0 m/s. The second is thrown downward, at a velocity of −12.0 m/s. Ignore air resistance. Determine (a) how long it takes the first rock to hit the ground and (b) at what velocity it hits. Determine (c) how long it takes the second rock to hit the ground and (b) at what velocity it hits.

(a) __________ s
(b) __________ m/s
(c) __________ s
(d) __________ m/s

23.11 A person throws a ball straight up. He releases the ball at a height of 1.75 m above the ground and with a velocity of 12.0 m/s. Ignore the effects of air resistance. (a) How long until the ball reaches its highest point? (b) How high above the ground does the ball go?

(a) __________ s
(b) __________ m

23.12 To determine how high a cliff is, a llama farmer drops a rock, and then 0.800 s later, throws another rock straight down at a velocity of −10.0 m/s. Both rocks land at the same time. How high is the cliff?

__________ m

23.13 Suppose you drop a ball onto the ground (starting it with zero initial velocity). The ball bounces back upward with half the speed at which it hit the ground. How high up does the ball bounce back, as a percentage of the initial height?

__________ percent

Section 25 - Interactive problem: soccer on the Moon

25.1 Use the information given in the interactive problem in this section to answer the following questions. (a) What is the initial velocity of the ball needed to match the graph? (b) What is the acceleration? Give both answers with two significant figures. Test your answer using the simulation.

(a) __________ m/s
(b) __________ m/s^2

Section 27 - Spreadsheet: modeling air resistance

27.1 For an object moving at certain speeds in certain fluids, the acceleration of the object due to the resistance of the fluid is proportional to the square of its velocity. For a golf ball in water, this acceleration is kv^2 where $k = -18.7\ \text{m}^{-1}$. A golf ball enters a container of water moving at a velocity of 10.0 m/s and all of its acceleration is due to the resistance of the fluid. In other words, ignore the force of gravity. How long does it take for the ball to be moving at 10.0% of its initial speed?

__________ s

Section 29 - Interactive problem: shuffleboard

29.1 Use the information given in the interactive problem in this section to find the initial velocity required to stop the puck at the desired location. Test your answer using the simulation.

__________ m/s

29.2 Use the information given for the second interactive problem in this section to answer the following questions. (a) Using an initial velocity of 2.00 m/s, what is the acceleration of the puck? (b) What is the initial velocity required to stop the puck at the goal line? Test your answer using the simulation.

(a) __________ m/s^2
(b) __________ m/s

Additional Problems

A.1 A cheetah is traveling across the African countryside, initially moving at a constant velocity of 9.0 m/s. It sees a gazelle at t = 0 seconds and begins to accelerate in a straight line. At t = 4.0 seconds it has a velocity of 23 m/s and then continues to move at this constant velocity. At t = 5.0 seconds, the cheetah is standing still, having knocked over the gazelle. (a) What is the average acceleration of the cheetah from t = 0 to t = 4.0 seconds? (b) What is the average acceleration of the cheetah from t = 0 to t = 5.0 seconds?

(a) __________ m/s^2
(b) __________ m/s^2

A.2 On Deimos (a moon of Mars) the free fall acceleration is 3.0×10^{-3} m/s^2. A rock is thrown straight up from the surface of Deimos with an initial velocity of 0.30 m/s. (a) Plot a graph of the rock's vertical displacement versus time (displacement on the vertical axis, time on the horizontal axis) for about the first 200 seconds of its travel. (b) What geometric shape have you just plotted?

(a) Submit answer on paper.
(b) i. Parabola
ii. Hyperbola
iii. Semicircle
iv. Isosceles Triangle

A.3 Above the surface of the Earth, the magnitude of the acceleration due to gravity is inversely proportional to the square of the distance from the Earth's center. For objects falling to Earth, $a = k/r^2$, where k is a constant and r is the distance to the center of the Earth. (a) The radius of the Earth is about 6.40×10^3 km. Find k using the fact that $a = g$ at Earth's surface. (b) Find the magnitude of the acceleration due to Earth's gravity at 1.00×10^2 km above the surface of the Earth.

(a) __________ m^3/s^2
(b) __________ m/s^2

A.4 Rock climbers seek safety by placing "protection" in the rock. These are fixed pieces of equipment called "cams" or "friends" that their rope passes through. A climber's partner holds the other end of the rope, so if she falls, the protection catches and stops her fall. In some situations, a climber falls far enough that the force exerted on the protection pulls it out of the rock face. The climber then falls farther, and exerts even more force on the next piece of protection, and pulls it out too. This terrifying phenomenon is called "zippering a pitch."
Imagine a climber has climbed 5.00 m directly above her uppermost piece of protection, with no slack (excess length) in the rope, and the next piece of protection is 6.00 m directly below the uppermost piece of protection. She falls with an acceleration of 9.80 m/s^2 downward. The first piece of protection cuts her velocity in half, but then it fails and she falls farther. What is her speed just when the rope pulls on the second piece of protection?

__________ m/s

A.5 A kangaroo-like animal on a mysterious planet jumps and then moves in freefall motion, first straight up, and then straight down as the planet's gravitational acceleration reverses its direction. Its vertical position at a certain time is described by the function $y(t) = -t^2 + 4.0t + 4.0$, where y is in meters and t is in seconds. (a) Draw the position versus time graph from t = −2.0 s to t = 6.0 s. (b) Determine the displacement from t = 0 s to t = 2.0 s. (c) Determine the average vertical velocity from t = 0 s to t = 2.0 s. (d) Determine the average vertical velocity from t = 1.0 s to t = 2.0 s. (e) Draw a line tangent to the graph at t = 2.0 s. What is its slope? (f) Assume there is no air resistance. What is the acceleration due to gravity on this planet?

(a) Submit answer on paper.
(b) __________ m
(c) __________ m/s
(d) __________ m/s
(e) __________ m/s
(f) __________ m/s^2

A.6 Pyrotechnics experts are setting up rockets as part of a fireworks show. They have a standard fuel that provides a constant upwards acceleration for 5.00 seconds, and then the rockets are in free fall, accelerating at −9.80 m/s^2. The peak height of the rockets is exactly 160 meters. (a) Find the upwards acceleration of the rocket during the five-second period. (b) A blue rocket is launched at t = 0 s and is detonated by radio at its peak height. When is it detonated? (c) The rockets do not need to be at the peak of their trajectory to detonate. The experts want a red rocket and a white rocket to detonate simultaneously on their way up. The white rocket is to detonate at 110 meters and the red rocket, launched after it, is to detonate at 60.0 m. How much later than the white rocket should the red rocket be launched?

(a) __________ m/s^2
(b) __________ s
(c) __________ s

A.7 Lucy is using a programmable treadmill. She inputs a top speed into the treadmill and presses "start." The treadmill accelerates from rest at 0.100 m/s^2 until it reaches that speed, after which it stays at that constant speed. Later, when she is getting tired, she hits the "stop" button and the treadmill slows down at a rate of 0.200 m/s^2 until it comes to a stop. The treadmill tells her that she traveled the equivalent of 3.00 kilometers in 15.0 minutes while the treadmill was in motion. (a) If Lucy had been running forward in a straight line on the ground (instead of in place on the treadmill), what would be her average velocity in meters per second over the whole 15.0 minutes? (b) What would be her average acceleration over the 15.0 minutes? (c) What was the top speed of the treadmill? (d) How long after pressing "start" did she press "stop?"

(a) _________ m/s
(b) _________ m/s^2
(c) _________ m/s
(d) _________ s

A.8 Starting from rest at the ground floor, an elevator takes 19.0 seconds for a vertical trip of 70.0 meters. It begins its journey with 4.50 seconds of constant acceleration, then moves for 9.00 seconds at constant velocity, and finally, moves for 5.50 seconds of constant negative acceleration that brings the elevator to a stop. (a) What is the maximum velocity of the elevator? (b) What is the elevator's initial acceleration? (c) What is its final acceleration?

(a) _________ m/s
(b) _________ m/s^2
(c) _________ m/s^2

3.0 - Introduction

interactive 1

Use vectors to visit three space attractions

Knowing “how far” or “how fast” can often be useful, but “which way” sometimes proves even more valuable. If you have ever been lost, you understand that direction can be the most important thing to know.

Vectors describe “how much” **and** “which way,” or, in the terminology of physics, magnitude and direction. You use vectors frequently, even if you are not familiar with the term. “Go three miles northeast” or “walk two blocks north, one block east” are both vector descriptions. Vectors prove crucial in much of physics. For example, if you throw a ball up into the air, you need to understand that the initial velocity of the ball points “up” while the acceleration due to the force of gravity points “down.”

In this chapter, you will learn the fundamentals of vectors: how to write them and how to combine them using operations such as addition and subtraction.

On the right, a simulation lets you explore vectors, in this case displacement vectors. In the simulation, you are the pilot of a small spaceship. There are three locations nearby that you want to visit: a refueling station, a diner, and the local gym. To reach any of these locations, you describe the displacement vector of the spaceship by setting its x (horizontal) and y (vertical) components. In other words, you set how far horizontally you want to travel, and how far vertically. This is a common way to express a two-dimensional vector.

There is a grid on the drawing to help you determine these values. You, and each of the places you want to visit, are at the intersection of two grid lines. Each square on the grid is one kilometer across in each direction. Enter the values, press GO, and the simulation will show you traveling in a straight line − along the displacement vector − according to the values you set. See if you can reach all three places. You can do this by entering displacement values to the nearest kilometer, like (3, 4) km. To start over at any time, press RESET.

3.1 - Scalars

Scalar: A quantity that states only an amount.

Scalar quantities state an amount: “how much” or “how many.” At the right is a picture of a dozen eggs. The quantity, a dozen, is a scalar. Unlike vectors, there is no direction associated with a scalar − no up or down, no left or right − just one quantity, the amount. A scalar is described by a single number, together with the appropriate units.

Temperature provides another example of a scalar quantity; it gets warmer and colder, but at any particular time and place there is no “direction” to temperature, only a value. Time is another commonly used scalar.

Speed and distance are yet other scalars. A speed like 60 kilometers per hour says how fast but not which way. Distance is a scalar since it tells you how far away something is, but not the direction.

concept 1

Scalars
Amount
Only one value

concept 2

Examples of scalars
12 eggs
Temperature is –5º C

Vector: A quantity specified by both magnitude *and* direction.


A spelunker (cave explorer) uses both distance and direction to navigate.

Vectors have both magnitude (how much) **and** direction. For example, vectors can be used to supply traveling instructions. If a pilot is told “Fly 20 kilometers due south,” she is being given a displacement vector to follow. Its magnitude is 20 kilometers and its direction is south. Vector magnitudes are positive or zero; it would be confusing to tell somebody to drive negative 20 kilometers south.

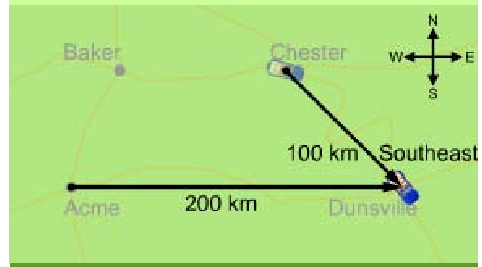
Many of the fundamental quantities in physics are vectors. For instance, displacement, velocity and acceleration are all vector quantities. Physicists depict vectors with arrows. The length of the arrow is proportional to the vector’s magnitude, and the arrow points in the direction of the vector. The horizontal vector in Concept 1 on the right represents the displacement of a car driving from Acme to Dunsville.

You see two displacement vectors in Concept 1. The displacement vector of a drive from Acme to Dunsville is twice as long as the displacement vector from Chester to Dunsville. This is because the distance from Acme to Dunsville is twice that of Chester to Dunsville.

Even if they do not begin at the same point, two vectors are equal if they have the same magnitude and direction. For instance, the vector from Chester to Dunsville in Concept 1 represents a displacement of 100 km southeast. That vector could be moved without changing its meaning. Perhaps it is 100 km southeast from Edwards to Frankville, as well. A vector’s meaning is defined by its length and direction, not by its starting point.

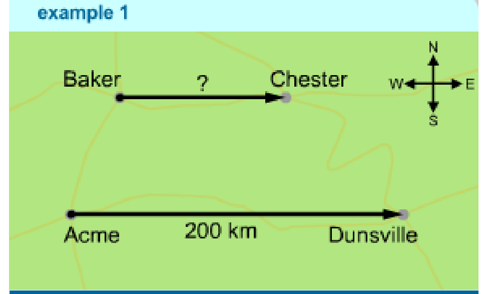
Now that we have introduced the concept of vectors formally, we will express vector quantities in boldface. For instance, $\mathbf{F}$ represents force, $\mathbf{v}$ stands for velocity, and so on. You will often see F and v, as well, representing the magnitudes of the vectors, without boldface. Why? Because it is frequently useful to discuss the magnitude of the force or the velocity without concerning ourselves with its direction. For instance, there may be several equations that determine the magnitude of a vector quantity like force, but not its direction.

concept 1



Vectors
Magnitude and direction
Represented by arrows
Length proportional to magnitude

example 1



It is half as far from Baker to Chester as from Acme to Dunsville. Describe the displacement vector from Baker to Chester.
Displacement: 100 km, east

3.3 - Polar notation

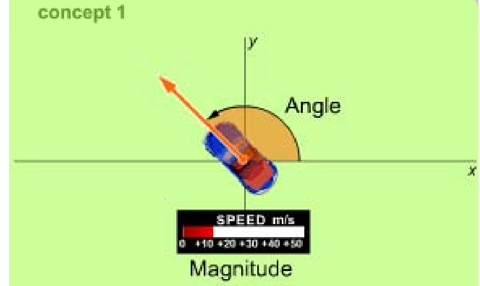
Polar notation: Defining a vector by its angle and magnitude.

Polar notation is a way to specify a vector. With polar notation, the magnitude and direction of the vector are stated separately. Three kilometers due north is an example of polar notation. “Three kilometers” is the magnitude and “north” is the direction. The magnitude is always stated as a positive value. Instead of using “compass” or map directions, physicists use angles. Rather than saying “three kilometers north,” a physicist would likely say “three kilometers directed at 90 degrees.”

The angle is most conveniently measured by placing the vector’s starting point at the origin. The angle is then typically measured from the positive side of the x axis to the vector. This is shown in Concept 1 to the right.

Angles can be positive or negative. A positive angle indicates a counterclockwise

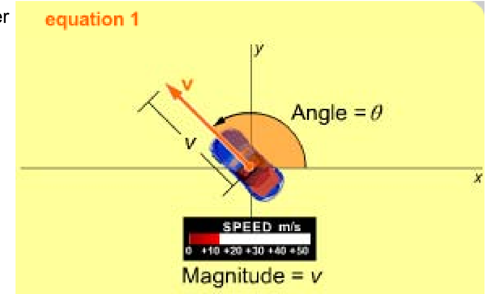
concept 1



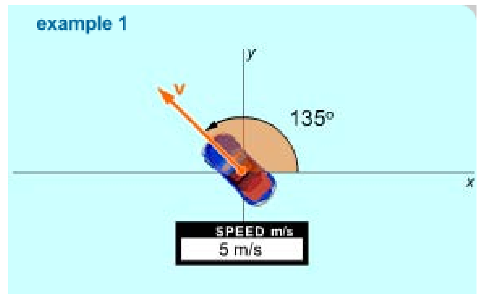
Polar notation
Magnitude and angle

direction, a negative angle a clockwise direction. For example, $90°$ represents a quarter turn **counterclockwise** from the positive x axis. In other words, a vector with a $90°$ angle points straight up. We could also specify this angle as $-270°$.

The radian is another unit of measurement for angles that you may have seen before. We will use degrees to specify angles unless we specifically note that we are using radians. (Radians do prove essential at times.)



Polar notation
v is magnitude
θ is angle
Written $\mathbf{v} = (v, \theta)$



Write the velocity vector of the car in polar notation.
$\mathbf{v} = (v, \theta)$
$\mathbf{v} = (5\text{ m/s}, 135°)$

3.4 - Vector components and rectangular notation

Rectangular notation: Defining a vector by its components.

Often what we know, or want to know, about a particular vector is not its overall magnitude and direction, but how far it extends horizontally and vertically. On a graph, we represent the horizontal direction as x and the vertical direction as y. These are called *Cartesian coordinates*. The x *component* of a vector indicates its extent in the horizontal dimension and the y *component* its extent in the vertical dimension.

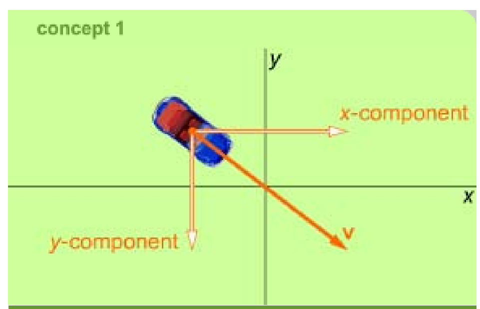
Rectangular notation is a way to describe a vector using the components that make up the vector. In rectangular notation, the x and y components of a vector are written inside parentheses. A vector that extends a units along the x axis and b units along the y axis is written as (a, b). For instance (3, 4) is a vector that extends positive three in the x direction and positive four in the y direction from its starting point.

The components of vectors are scalars with the direction indicated by their sign: x components point right (positive) or left (negative), and y components point up (positive) or down (negative). You see the x and y components of a car's velocity vector in Concept 1 at the right, shown as "hollow" vectors. The x and y values define the vector, as they provide direction and magnitude.

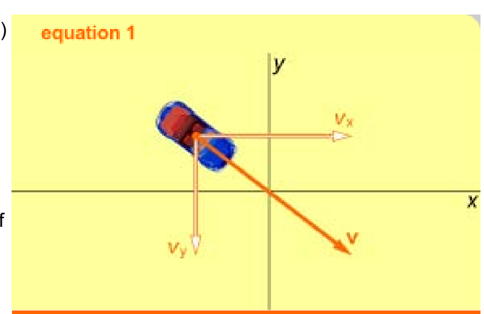
For a vector $\mathbf{A}$, the x and y components are sometimes written as A_x and A_y. You see this notation used for a velocity vector $\mathbf{v}$ in Equation 1 and Example 1 on the right.

Consider the car shown in Example 1 on the right. Its velocity has an x component v_x of 17 m/s and a y component v_y of −13 m/s. We can write the car's velocity vector as (17, −13) m/s.

A vector can extend in more than two dimensions: z represents the third dimension. Sometimes z is used to represent distance toward or away from you. For instance, your computer monitor's width is measured in the x dimension, its height with y and your distance from the monitor with z. If you are reading this on a computer monitor and punch your computer screen, your fist would be moving in the z dimension. (We hope



Vector components and rectangular notation
x component and y component



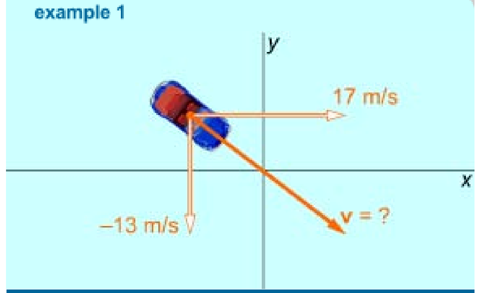
Rectangular notation
v_x is horizontal component

we're not the cause of any such aggressive feelings.) Three-dimensional vectors are written as (x, y, z).

The z component can also represent altitude. A Tour de France bike racer might believe the z dimension to be the most important as he ascends one of the competition's famous climbs of a mountain pass.

v_y is vertical component
Written $\mathbf{v} = (v_x, v_y)$

example 1



What is the car's velocity vector in rectangular notation?

$\mathbf{v} = (v_x, v_y)$ m/s

$\mathbf{v} = (17, -13)$ m/s

3.5 - Adding and subtracting vectors graphically

Vectors can be added and subtracted. In this section, we show how to do these operations graphically. For instance, consider the vectors $\mathbf{A}$ and $\mathbf{B}$ shown in Concept 1 to the right. The vector labeled $\mathbf{A} + \mathbf{B}$ is the sum of these two vectors.

It may be helpful to imagine that these two vectors represent displacement. A person walks along displacement vector $\mathbf{A}$ and then along displacement vector $\mathbf{B}$. Her initial point is the origin, and she would end up at the point at the end of the $\mathbf{A} + \mathbf{B}$ vector. The sum represents the displacement vector from her initial to final position.

To be more specific about the addition process: We start with two vectors, $\mathbf{A}$ and $\mathbf{B}$, both drawn starting at the origin (0, 0). To add them, we move the vector $\mathbf{B}$ so it starts at the head of $\mathbf{A}$. The diagram for Equation 1 shows how the $\mathbf{B}$ vector has been moved so it starts at the head of $\mathbf{A}$. The sum is a vector that starts at the tail of $\mathbf{A}$ and ends at the head of $\mathbf{B}$.

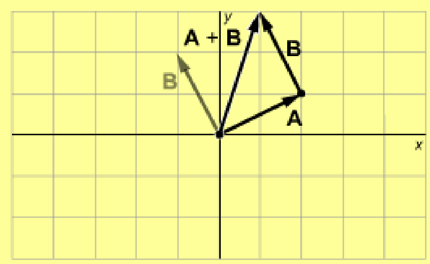
In summary, to add two vectors, you:

1. Place the tail of the second vector at the head of the first vector. (The order of addition does not matter, so you can place the tail of the first vector at the head of the second as well.)
2. Draw a vector between the tail end of the first vector and the head of the second vector. This vector represents the sum of two vectors.

To emphasize a point: You can think of this as combining a series of vector instructions. If someone says, "Walk positive three in the x direction and then negative two in the y direction," you follow one instruction and then the other. This is the equivalent of placing one vector's tail at the head of the other. An arrow from where you started to where you ended represents the resulting vector. Any vector is the vector sum of its rectangular components.

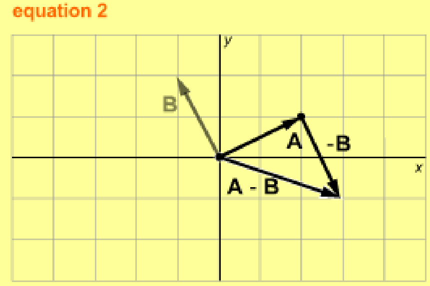
When two vectors are parallel and pointing in the same direction, adding them is relatively simple: You just combine the two arrows to form a longer arrow. If the vectors are parallel but pointing in opposite directions, the result is a shorter arrow (three steps forward plus two steps back equals one step forward).

equation 1



Adding vectors A + B graphically
Move tail of B to head of A
Draw vector from tail of A to head of B

equation 2



Subtracting A – B graphically
Take the opposite of B
Move it to head of A
Draw vector from tail of A to head of –B

To subtract two vectors, take the opposite of the vector that is being subtracted, and then add. (The opposite or negative of a vector is a vector with the same magnitude but opposite direction.) This is the same as scalar subtraction (for example 20 − 5 is the same as 20 + (−5)). To draw the opposite of a vector, draw it with the same length but the opposite direction. In other words, it starts at the same point but is rotated $180°$. The diagram for Equation 2 shows the subtraction of two vectors.

When a vector is added to its opposite, the result is the *zero vector*, which has zero magnitude and no direction. This is analogous to adding a scalar number to its opposite, like adding +2 and −2 to get zero.

3.6 - Adding and subtracting vectors by components

You can combine vectors graphically, but it may be more precise to add up their components.

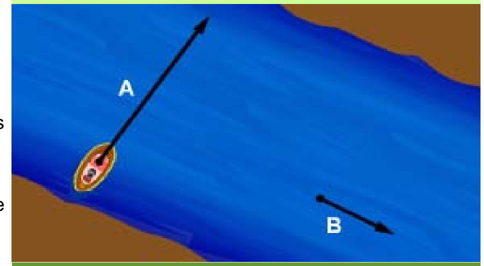
You perform this operation intuitively outside physics. If you were a dancer or a cheerleader, you would easily understand the following choreography: "Take two steps forward, four steps to the right and one step back." These are vector instructions. You can add them to determine the overall result. If asked how far **forward** you are after this dance move, you would say "one step," which is two steps forward plus one step back. You realize that your progress forward or back is unaffected by steps to the left or right. You correctly process left/right and forward/back separately. If a physics-oriented dance instructor asked you to describe the results of your "dancing vector" math, you would say, "One step forward, four steps to the right."

You have just learned the basics of vector addition, which is reasonably straightforward: Break the vector into its components and add each component independently. In physics though, you concern yourself with more than dance steps. You might want to add the vector (20, −40, 60) to (10, 50, 10). Let's assume the units for both vectors are meters. As with the dance example, each component is added independently. You add the first number in each set of parentheses: 20 plus 10 equals 30, so the sum along the x axis is 30. Then you add −40 and 50 for a total of 10 along the y axis. The sum along the z axis is 60 plus 10, or 70. The vector sum is (30, 10, 70) meters. If following all this in the text is hard, you can see another problem worked in Example 1 on the right.

Although we use displacement vectors in much of this discussion since they may be the most intuitive to understand, it is important to note that all types of vectors can be added or subtracted. You can add two velocity vectors, two acceleration vectors, two force vectors and so on. As illustrated in the example problem, where two velocity vectors are added, the process is identical for any type of vector.

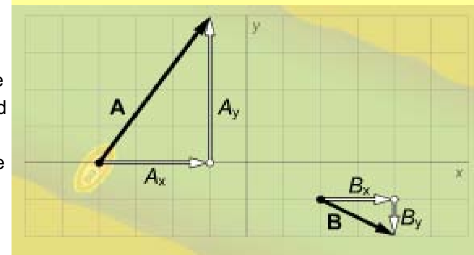
Vector subtraction works similarly to addition when you use components. For example, (5, 3) minus (2, 1) equals 5 minus 2, and 3 minus 1; the result is the vector (3, 2).

concept 1



Adding and subtracting vectors by components
Add (or subtract) each component separately

equation 1

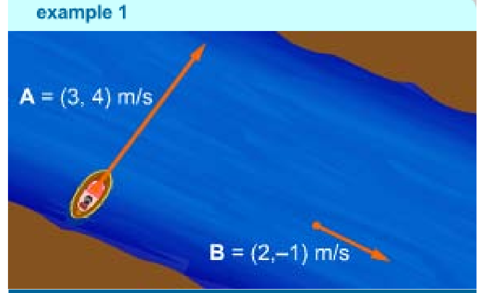


Adding and subtracting vectors by components

$$\mathbf{A} + \mathbf{B} = (A_x + B_x, A_y + B_y)$$
$$\mathbf{A} - \mathbf{B} = (A_x - B_x, A_y - B_y)$$

$\mathbf{A}$, $\mathbf{B}$ = vectors
A_x, A_y = $\mathbf{A}$ components
B_x, B_y = $\mathbf{B}$ components

example 1



The boat has the velocity A in still water. Calculate its velocity as the sum of A and the velocity B of the river's current.

$\mathbf{v} = \mathbf{A} + \mathbf{B}$
$\mathbf{v} = (3, 4)$ m/s + $(2, -1)$ m/s
$\mathbf{v} = (3 + 2, 4 + (-1))$ m/s
$\mathbf{v} = (5, 3)$ m/s

3.7 - Interactive checkpoint: vector addition

$$\mathbf{v}_1 = (5, 6)$$
$$\mathbf{v}_2 = (a, 4)$$
$$\mathbf{v}_1 + \mathbf{v}_2 = (2, b)$$

What are the number values of the constants a and b?

Answer:

a = ______ , b = ______

3.8 - Interactive checkpoint: a jogger

A jogger breaks her workout into three segments: jogging, sprinting and walking. Starting at home, she jogs a displacement vector of (a, 2a) blocks, sprints a displacement of (3b, b) blocks, and walks back home with a displacement of (2, −6) blocks. What is the vector value of her displacement during the sprint?

Answer:

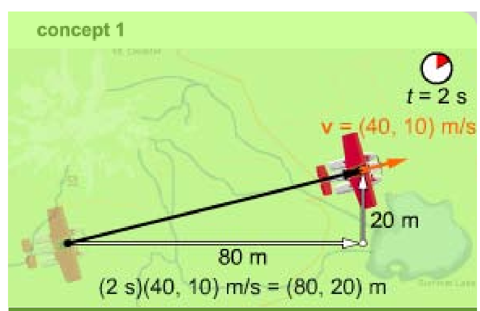
S = (______ , ______) blocks

3.9 - Multiplying rectangular vectors by a scalar

You can multiply vector quantities by scalar quantities. Let's say an airplane, as shown in Concept 1 on the right, travels at a constant velocity represented by the vector (40, 10) m/s. Let's say you know its current position and want to know where it will be if it travels for two seconds. Time is a scalar. To calculate the displacement, multiply the velocity vector by the time.

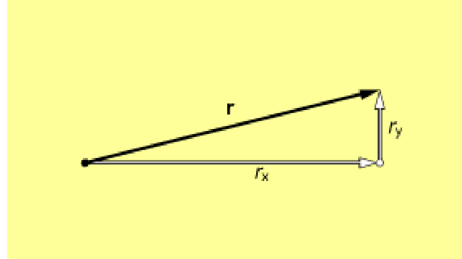
To multiply a vector by a scalar, multiply each component of the vector by the scalar. In this example, (2 s)(40, 10) m/s = (80, 20) m. This is the plane's displacement vector after two seconds of travel.

If you wanted the opposite of this vector, you would multiply by negative one. The result in this case would be (−40, −10) m/s, representing travel at the same speed, but in the opposite direction.



Multiplying a rectangular vector by a scalar
Multiply each component by scalar
Positive scalar does not affect direction

equation 1



Multiplying a rectangular vector by a scalar

$$s\mathbf{r} = (sr_x, sr_y)$$

s = a scalar
$\mathbf{r}$ = a vector
r_x, r_y = $\mathbf{r}$ components

example 1

What is the displacement d of the plane after 5.0 seconds?

$\mathbf{d}$ = (5.0 s)$\mathbf{v}$
$\mathbf{d}$ = (5.0 s) (12 m/s, 15 m/s)
$\mathbf{d}$ = ((5.0 s)(12 m/s), (5.0 s)(15 m/s))
$\mathbf{d}$ = (60, 75) m

3.10 - Multiplying polar vectors by a scalar

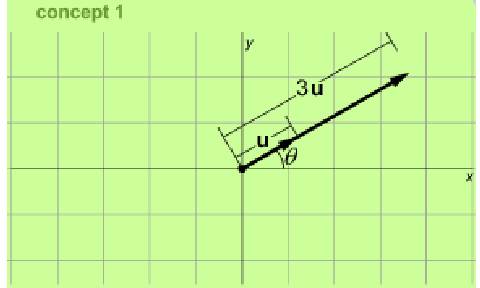
Multiplying a vector represented in polar notation by a positive scalar requires only one multiplication operation: Multiply the magnitude of the vector by the scalar. The angle is unchanged.

Let's say there is a vector of magnitude 50 km with an angle of 30°. You are asked to multiply it by positive three. This situation is shown in Example 1 to the right. Since you are multiplying by a positive scalar, the angle stays the same at 30°, and so the answer is 150 km at 30°.

If you multiply a vector by a negative scalar, multiply its magnitude by the absolute value of the scalar (that is, ignore the negative sign). Then change the direction of the vector by 180° so that it points in the opposite direction. In polar notation, since the magnitude is always positive, you add 180° to the vector's angle to take its opposite. The result of multiplying (50 km, 30°) by negative three is (150 km, **210**°).

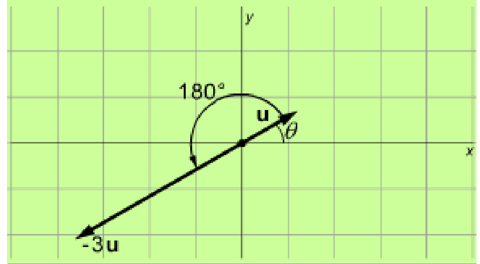
If adding 180° would result in an angle greater than 360°, then subtract 180° instead. For instance, in reversing an angle of 300°, subtract 180° and express the result as 120° rather than 480°. The two results are identical, but 120° is easier to understand.

concept 1



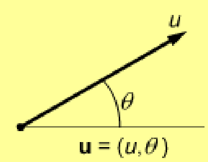
Multiplying polar vector by positive scalar
Multiply vector's magnitude by scalar
Angle unchanged

concept 2



Multiplying by negative scalar
Use absolute value and reverse direction

equation 1



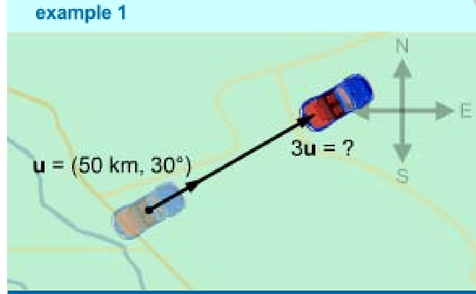
Multiplying by negative scalar

$s\mathbf{u} = (su, \theta)$, if s positive

$s\mathbf{u} = (|s|u, \theta + 180°)$, if s negative

s = a scalar, $\mathbf{u}$ = a vector
u = magnitude of vector
θ = angle of vector

example 1



What is the displacement vector if the car travels three times as far?

$s\mathbf{u} = (su, \theta)$

$3\mathbf{u} = (\ 3(50 \text{ km}), 30°)$

$3\mathbf{u} = (150 \text{ km}, 30°)$

3.11 - Converting vectors from polar to rectangular notation

You may find it useful at times to convert a vector expressed in polar notation to rectangular coordinates. To illustrate, what if someone gave you these directions: "Travel 3.0 km at $35°$ and then 2.0 km at $-15°$." You suspect you could shorten this trip by adding these two vectors and just traveling the resultant vector, but how would you add them? To add them algebraically (as opposed to graphically), it is simpler if you first convert both to rectangular vectors.

Converting the vectors above requires some trigonometry basics, namely sines and cosines. In short, you treat the magnitude of a vector as the hypotenuse of a right triangle, with the x component as its horizontal leg and the y component as its vertical leg.

If you want to convert the first vector above, take 3.0 km as the hypotenuse. Then calculate the x component by multiplying 3.0 km by cos $35°$, and the y component by multiplying 3.0 km by sin $35°$. Here, $x = (3.0\text{ km})(0.82)$ and $y = (3.0\text{ km})(0.57)$, so the vector in rectangular coordinates is (2.5, 1.7) km.

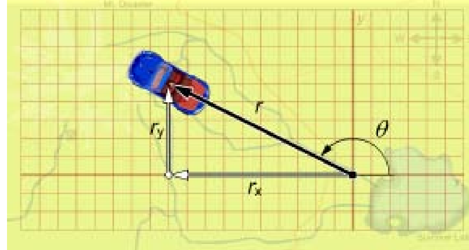
Using the same method with the other vector, 2.0 km at $-15°$ equals (1.9, −0.52) km. The positive x component and negative y component indicate that this vector points down and to the right, the correct direction for a vector with an angle of $-15°$.

We began this section by asking you how you would add these two vectors. Our work has made this an easier problem: (2.5, 1.7) plus (1.9, −0.52) equals (4.4, 1.2). The units are kilometers.

The x and y components can be positive or negative. For instance, the x component will be negative when the cosine is negative, which it is for angles between $90°$ and $270°$. This corresponds to vectors that have an x component which points to the left. The y component will be positive when the sine is positive (between $0°$ and $180°$, the vector has an upward y component) and negative when the sine is negative (between $180°$ and $360°$, the vector has a downward y component).

Since it is easy to err, it is a good practice to compare directions and the signs of the components. In Example 1, the negative x component is correct, since the car is moving to the left. If we had calculated a negative y component, we have erred in our calculations, since the car is clearly moving “up” in the positive y direction.

equation 1



Converting a vector from polar to rectangular notation

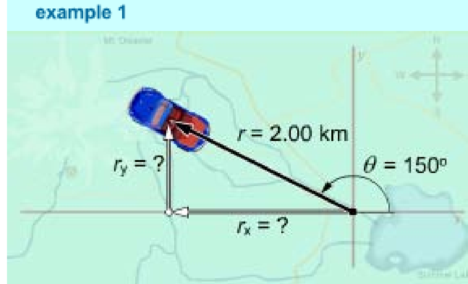
To express (r, θ) as (r_x, r_y)

$$r_x = r\cos\theta$$
$$r_y = r\sin\theta$$

r = magnitude, θ = angle

r_x, r_x = components of vector

example 1



What is the displacement vector r of the car in rectangular notation?

$r_x = r\cos\theta = (2.00\text{ km})(\cos 150º)$

$r_x = -1.73$ km

$r_y = r\sin\theta = (2.00\text{ km})(\sin 150º)$

$r_y = 1.00$ km

$\mathbf{r} = (x, y) = (-1.73, 1.00)$ km

3.12 - Converting vectors from rectangular to polar notation

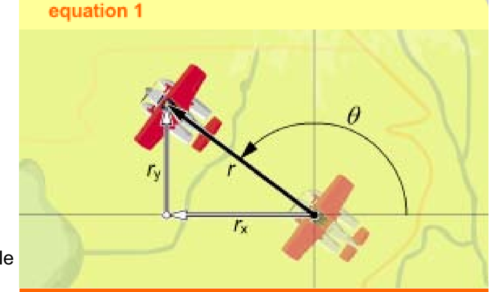
In some counties of the United States, the main roads travel either east-west or north-south. If you wanted to drive from one town to another, the roads might force you to travel 40 km west and then 30 km north. On the other hand, if you had a plane, you could fly in a straight line between the two towns, which would be a shorter distance. You would need to know the angle at which to fly and the distance. We work this problem out in Example 1, but before that, we review the concepts necessary to solve the problem.

To determine the angle and distance, you need to convert from rectangular to polar coordinates. You would use trigonometry to do so. The x and y components represent the legs of a triangle. You need to determine the length of the hypotenuse and the angle the hypotenuse makes with the positive x axis.

In Equation 1 on the right, you see that the Pythagorean theorem is used to calculate the hypotenuse when the two legs are known. The magnitude of the vector (the hypotenuse) is represented with r, and the two legs, called r_x and r_y here, are the components of the vector. In the example, the distance in kilometers is the square root of $(-40\text{ km})^2 + (30\text{ km})^2$. That works out to 50 km.

Now you can determine the angle, which we represent as θ. As you may recall, the tangent function relates the base and height of a right triangle to the angle between the hypotenuse and the base (in this case, the x axis). The angle θ is the arctangent of the

equation 1



Converting rectangular to polar

To express (r_x, r_y) as (r, θ)

ratio of the two legs of the triangle. You see this in Equation 1 also.

You need to be particularly careful with tangents and arctangents since, as Equation 2 shows, two different angles can have the same tangent. For instance, the vectors (1, 1) and (−1, −1) point in opposite directions, but since the ratio of their components in each case is 1, the tangent equals 1 in both cases.

To find the direction of our plane's travel, we need to know the arctangent of (30 km)/(−40 km), the ratio of the legs of the triangle. Our calculator reported back a value of about $-37°$. (Make sure you know if your calculator is set for degrees or radians!)

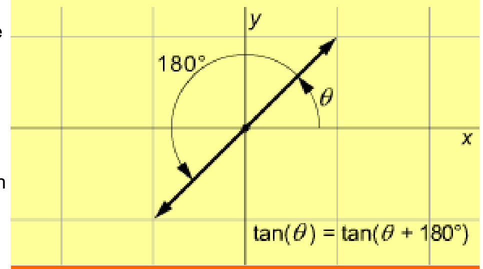
Unfortunately, this answer does not make sense given the circumstances described above. A vector at an angle of $-37°$ would have a positive x component and a negative y component. However, the plane's displacement vector has the opposite characteristics: a negative x component and a positive y component.

As mentioned earlier, two angles can generate the same value for a tangent, so we need to find the other angle. With arctangents, the "other angle" is found by adding or subtracting $180°$. Adding $180°$ to $-37°$ equals $143°$, and $143°$ describes a vector with a negative x value and a positive y value, which is appropriate for this situation.

$r = \sqrt{r_x^2 + r_y^2}$

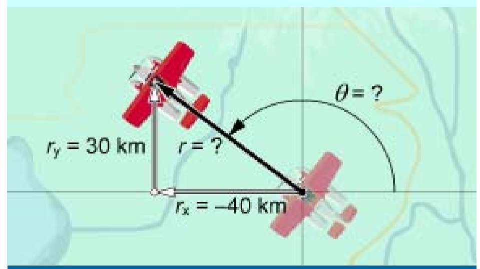
$\theta = \arctan(r_y/r_x)$ r_x, r_y = components of vector r = magnitude, θ = angle

equation 2



Converting rectangular to polar
Be sure to check for the correct angle/quadrant

example 1



What is the plane's displacement r in polar notation?

$r = \sqrt{r_x^2 + r_y^2}$

$r = \sqrt{(-40\ \text{km})^2 + (30\ \text{km})^2}$

$r = 50$ km

$\theta = \arctan(r_y/r_x)$

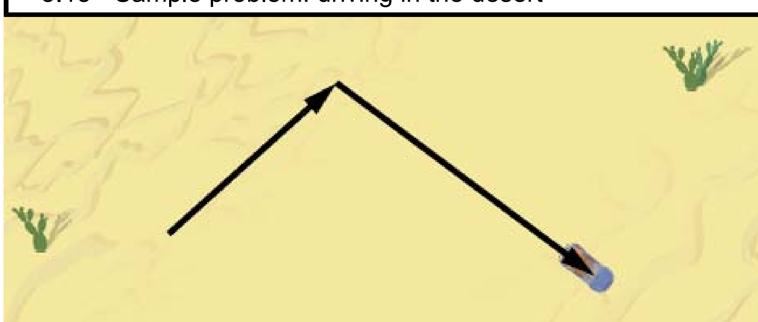
$\theta = \arctan(30\ \text{km}/–40\ \text{km})$

$\theta = \arctan(–0.75)$

$\theta = –37° + 180° = 143°$

$\mathbf{r} = (r, \theta) = (50\ \text{km}, 143°)$

3.13 - Sample problem: driving in the desert

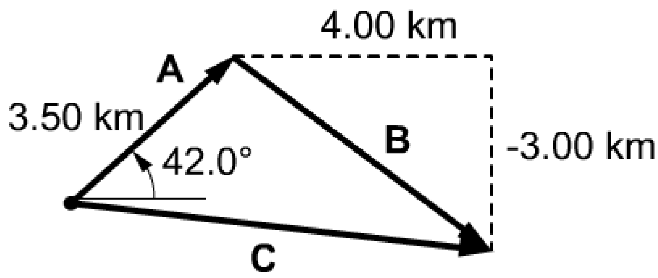


You are told to drive 3.50 km at 42.0°, then drive as directed by a vector of (4.00, −3.00) km.

What is your resulting displacement in rectangular coordinates? In polar notation?

You are given driving directions as two displacement vectors, one stated with polar values and the second with rectangular components. You are asked to find the resulting displacement vector in both rectangular and polar notation.

Draw a diagram



Variables

We use $\mathbf{A}$ to indicate the first vector, $\mathbf{B}$ for the second vector, and $\mathbf{C}$ for their sum.

	polar notation	rectangular notation
vector $\mathbf{A}$	(3.50 km, 42.0°)	(A_x, A_y)
vector $\mathbf{B}$	not needed	(4.00, −3.00) km
vector sum $\mathbf{C}$	(C, θ)	(C_x, C_y)

What is the strategy?

1. Convert the first vector $\mathbf{A}$ to rectangular notation.
2. Add vectors $\mathbf{A}$ and $\mathbf{B}$ by adding their components. This will give you the resulting displacement $\mathbf{C}$ in rectangular notation.
3. Convert $\mathbf{C}$ to polar notation. Check to make sure the angle is in the right quadrant.

Mathematics principles

Polar to rectangular conversion

$$r_x = r\cos\theta$$

$$r_y = r\sin\theta$$

Rectangular to polar conversion

$$r = \sqrt{r_x^2 + r_y^2}$$

$$\theta = \arctan(r_y/r_x)$$

Adding vectors

$$\mathbf{A} + \mathbf{B} = (A_x + B_x, A_y + B_y)$$

Step-by-step solution

We start by converting vector $\mathbf{A}$ to rectangular notation.

Step	Reason
1. $A_x = A\cos\theta$	x component of vector
2. $A_x = (3.50\text{ km})(\cos 42.0°)$	enter values
3. $A_x = 2.60\text{ km}$	cosine, multiplication
4. $A_y = A\sin\theta$	y component of vector
5. $A_y = (3.50\text{ km})(\sin 42.0°)$	enter values
6. $A_y = 2.34\text{ km}$	sine, multiplication
7. $\mathbf{A} = (2.60, 2.34)\text{ km}$	combine components

We do not need to convert **B** to rectangular notation, since it was given to us in that form. Now we add the vectors. This gives us the answer to the first question.

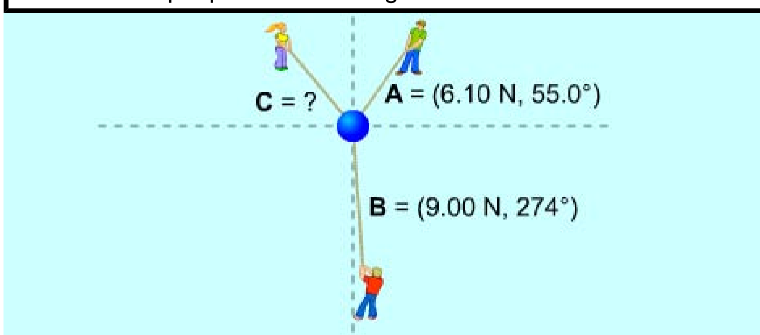
Step	Reason
8. $\mathbf{C} = \mathbf{A} + \mathbf{B}$	desired vector is sum
9. $\mathbf{C} = (A_x + B_x, A_y + B_y)$	sum of two vectors
10. $\mathbf{C} = (2.60 + 4.00, 2.34 + (–3.00))$	enter values
11. $\mathbf{C} = (6.60, –0.66)$ km	addition

Finally, we convert **C** to polar notation: magnitude and angle.

Step	Reason
12. $C = \sqrt{C_x^2 + C_y^2}$	equation for magnitude
13. $C = \sqrt{6.60^2 + (-0.66)^2}$	enter values
14. $C = 6.63$ km	evaluate
15. $\theta = \arctan (C_y/C_x)$	equation for angle
16. $\theta = \arctan (–0.66/6.60)$	enter values
17. $\theta = –5.71°$	evaluate
18. $\mathbf{C} = (6.60 \text{ km}, –5.71°)$	state magnitude and angle

Once we determined the signs of the x and y components of **C** in step 11, we knew **C** pointed down and to the right. This means a negative angle is appropriate for the polar notation. We could also have stated this angle as $354.32°$ ($= 360° - 5.68°$).

3.14 - Sample problem: looking ahead to forces



You pull on the ball with a force of 6.10 newtons at 55.0°. Your brother applies 9.00 newtons of force at 274°. The ball is not moving.

What is your sister's force vector in polar notation?

In a later chapter you will learn about the concept of a force. Force is a vector quantity that follows the laws of vectors that you have learned in this section. The unit of force is the newton (N).

If something is neither moving nor accelerating, it means that the vector sum of the forces acting on it is zero. This principle allows us to solve the problem, since it means the force vectors must sum to zero.

Variables

	polar notation	rectangular notation
your force, **A**	(6.10 N, 55.0°)	(A_x, A_y)
your brother's force, **B**	(9.00 N, 274°)	(B_x, B_y)
your sister's force, **C**	(C, θ)	(C_x, C_y)

What is the strategy?

1. Convert **A** and **B** to rectangular notation.
2. Set the sum of the x components of all three vectors to zero, and solve for the x component of **C**.
3. Set the sum of the y components to zero, and solve for the y component of **C**.
4. Convert **C** to polar notation.

Mathematics principles

Polar to rectangular conversion

$r_x = r\cos\theta$

$r_y = r\sin\theta$

Rectangular to polar conversion

$r = \sqrt{r_x^2 + r_y^2}$

$\theta = \arctan(r_y/r_x)$

Adding three vectors

$\mathbf{A} + \mathbf{B} + \mathbf{C} = (A_x + B_x + C_x, A_y + B_y + C_y)$

Step-by-step solution

First, we convert vector **A** to rectangular notation.

Step	Reason
1. $A_x = A\cos\theta$	x component of vector
2. $A_x = (6.10\text{ N})(\cos 55.0°)$	enter values
3. $A_x = 3.50\text{ N}$	evaluate
4. $A_y = A\sin\theta$	y component of vector
5. $A_y = (6.10\text{ N})(\sin 55.0°)$	enter values
6. $A_y = 5.00\text{ N}$	evaluate
7. $\mathbf{A} = (3.50, 5.00)\text{ N}$	combine components

We use steps similar to those above to convert **B** to rectangular notation.

Step	Reason
8. $B_x = (9.00\text{ N})(\cos 274°)$	enter values
9. $B_x = 0.628\text{ N}$	evaluate
10. $B_y = (9.00\text{ N})(\sin 274°)$	enter values
11. $B_y = -8.98\text{ N}$	evaluate
12. $\mathbf{B} = (0.628, -8.98)\text{ N}$	combine components

Now we add the x components of all three vectors and set the sum equal to zero. This lets us solve for the x component of vector **C**.

Step	Reason
13. $\mathbf{A} + \mathbf{B} + \mathbf{C} = \mathbf{0}$	vector sum is zero
14. $A_x + B_x + C_x = 0$	sum of the x components is zero
15. $3.50 + 0.628 + C_x = 0$	enter values
16. $C_x = -4.13\text{ N}$	solve for C_x

Similarly, we find the y component of **C**.

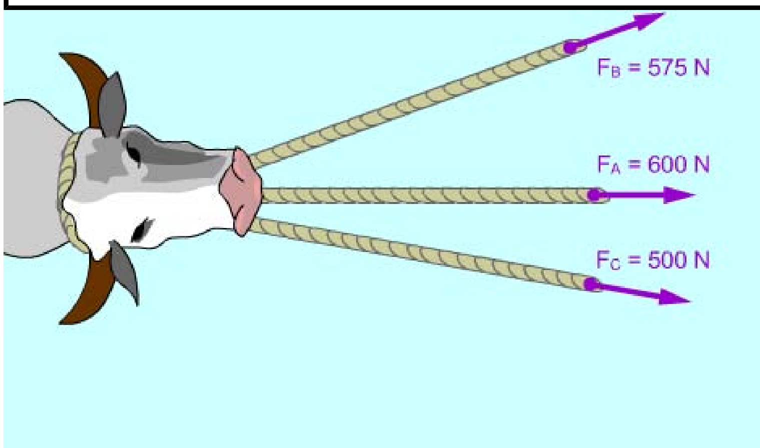
Step	Reason
17. $A_y + B_y + C_y = 0$	sum of the y components is zero
18. $5.00 + -8.98 + C_y = 0$	enter values
19. $C_y = 3.98\text{ N}$	solve for C_y

Finally, we convert this vector to polar notation: magnitude and angle.

Step	Reason
20. $C=\sqrt{C_x^2+C_y^2}$	equation for magnitude
21. $C=\sqrt{(-4.13)^2+(3.98)^2}$	enter values
22. $C = 5.74$ N	evaluate
23. $\theta = \arctan\,(C_y/C_x)$	equation for angle
24. $\theta = \arctan\,(3.98/-4.13)$	enter values
25. $\theta = -43.9°$	evaluate
26. $\theta = -43.9° + 180° = 136°$	add 180°
27. $\mathbf{C} = (5.74\text{ N}, 136°)$	combine magnitude and angle

Since **C** is your sister's force vector, she pulls with a force of 5.74 N in the direction 136°.

3.15 - Interactive checkpoint: a bum steer



Alvaro, a rancher, finds one of his steers trapped in quicksand. He tries to free the beast by attaching a rope and pulling with 600 newtons of horizontal force, to no avail. He asks his friends for help. Benjamin ties his rope to the steer, and pulls horizontally with 575 N at a 20.0° angle to Alvaro. Carlos does the same, and pulls on the other side of Alvaro at a −10.0° angle with 500 N of force. Alvaro's force is directed at 0°. What is the magnitude of their combined force on the steer?

Answer:

F = ______ N

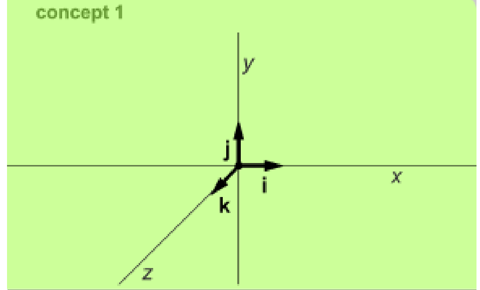
3.16 - Unit vectors

A vector can be described by its components in rectangular notation, as with (20, 30, 40). This describes a vector that extends 20 units in the x direction, 30 units in the y direction, and 40 units in the z direction. This form of notation is often used, but as your studies advance, you may use another form of notation called *unit vector* notation that has much in common with rectangular notation.

You see the general equation for unit vector notation in Equation 1 on the right. In this notation, the vector (a, b, c) is written as $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$. The unit vectors **i**, **j** and **k** have lengths equal to one and point along the x, y and z dimensions respectively. So, $a\mathbf{i}$ is the product of the scalar a and the unit vector **i**. If a is positive, the result is a vector of magnitude a pointing in the positive x direction. If a is negative, the vector $a\mathbf{i}$ points in the negative x direction, with magnitude equal to the absolute value of a.

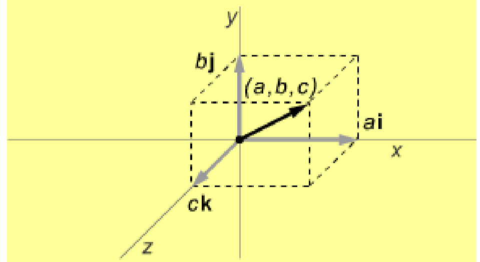
Unit vectors are dimensionless; there are no units associated with them. The product $a\mathbf{i}$ will have the same units as a. For example, $(3\text{ m/s})\mathbf{i} + (4\text{ m/s})\mathbf{j}$ represents a velocity vector of three meters per second in the x direction and four meters per second in the y direction. This can also be written as $3\mathbf{i} + 4\mathbf{j}$ m/s.

concept 1



Unit vectors
Represent dimensions (e.g. x, y, z)
Magnitude = one

equation 1

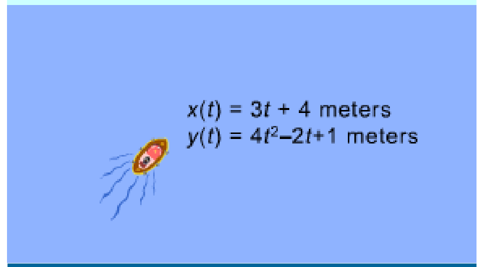


Unit vectors

$(a, b, c) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$

a, b, c = vector components
i, **j**, **k** = unit vectors

example 1



The x and y components of the boat's displacement are defined by the two equations shown. Express the displacement at 3.0 seconds as a vector r in unit vector notation.

$x(3.0) = 3(3.0) + 4 = 13$ m

$y(3.0) = 4(3.0)^2 - 2(3.0) + 1 = 31$ m

$\mathbf{r} = (13\mathbf{i} + 31\mathbf{j})$ m

3.17 - Interactive summary problem: back to base

In the simulation on the right, three spaceships need to reach their respective docking stations. (The red ship docks at the red station, the yellow ship at the yellow station and so on.) Your goal is to get them all home by calculating and entering the displacements from each of the ships to the corresponding stations.

For each of the ships you need to use a different way of describing the displacement vector. The red ship's displacement is specified with rectangular coordinates. The yellow ship must be specified with polar notation. And the purple ship's displacement is some scalar multiple of (2 km, 1 km). You need to calculate and enter the appropriate scalar value.

There is a grid on the drawing to help you determine the displacements. Each of the ships and docks is at the intersection of two grid lines. Each square on the grid is one kilometer across in each direction.

To dock the ships, first calculate the displacement for each ship in rectangular coordinates.

Since the red ship uses rectangular coordinates, enter those values to the nearest kilometer and you are done with that vessel.

interactive 1

Convert the yellow ship's displacement to polar notation. Enter the magnitude to the nearest 0.1 km and the angle to the nearest degree.

Determine what scalar multiple of (2, 1) will give you the purple ship's displacement, and enter it to the nearest whole integer.

After you enter the values, press GO, and see if the three ships arrive at their docking stations. Press RESET to start over. If you have difficulty, review the sections on converting rectangular to polar notation, and on multiplying vectors by scalars.

3.18 - Gotchas

Confusing the sine and cosine when converting from polar to rectangular notation. For a vector of magnitude r making an angle θ with the x axis, the x component is $r \cos \theta$ and the y component is $r \sin \theta$. It is easy to forget and incorrectly use $r \sin \theta$ for the x component or $r \cos \theta$ for the y component. There is a good check, used even by distinguished professors: Use $0°$ or $90°$ to verify that you are using the correct trigonometric function. For instance, if you thought the y component was proportional to $\cos \theta$, using $90°$ to check that assumption would indicate the error, because an angle of $90°$ corresponds to a vertical vector, but the cosine of $90°$ is zero.

Selecting the wrong value of arctan when converting from rectangular to polar notation. Pay attention to the quadrant where the vector lies. For example, (−4, −3) lies in the third quadrant, but arctan (−3/−4) = arctan (0.75), and your calculator will tell you that the value is $37°$. In this case you need to add $180°$ to the calculator result: The angle $217°$ has the same tangent and it lies in the correct quadrant.

In general, check that your answers indicate a vector that points in the direction you expect!

Stating a value as a scalar when a vector is required. This happens in physics and everyday life as well. You need to use a vector when direction is required. Throwing a ball up is different than throwing a ball down; taking highway I-5 south is different than taking I-5 north.

Vectors always start at the origin. No, they can start at any location.

3.19 - Summary

A scalar is a quantity, such as time, temperature, or speed, which indicates only amount.

A vector is a quantity, like velocity or displacement, which has both magnitude and direction. Vectors are represented by arrows that indicate their direction. The arrow's length is proportional to the vector's magnitude. Vectors are represented with **boldface** symbols, and their magnitudes are represented with *italic* symbols.

One way to represent a vector is with polar notation. The direction is indicated by the angle between the positive x axis and the vector (measured in the counterclockwise direction). For example, a vector pointing in the negative y direction would have a direction of $270°$ in polar notation. The magnitude is expressed separately. A polar vector is expressed in the form (r, θ) where r is the magnitude and θ is the direction angle.

Another way to represent a vector is by using rectangular notation. The vector's x and y components are expressed as an ordered pair of numbers (x, y). The components of a vector $\mathbf{A}$ are also written as A_x and A_y.

To add vectors graphically, place the tail of one on the head of the other, then draw a vector that goes from the free tail to the free head: The new vector is the sum. To subtract, first take the opposite of the vector being subtracted, then add. (The opposite of a vector has the same magnitude, but it points in the opposite direction.)

You can also add and subtract vectors by writing them in rectangular notation and adding or subtracting the x and y components separately to find the x and y components of the sum or difference vector.

Multiplying a vector by a scalar is done differently in polar and rectangular notation. For a polar vector, multiply the magnitude by the scalar. If the result is a negative magnitude, reverse the direction of the vector by adding (or subtracting) 180° from the angle. For a rectangular vector, multiply each component by the scalar. To convert from polar to rectangular notation, or from rectangular to polar, use the equations shown to the right.

Equations

Polar notation

$$\mathbf{v} = (v, \theta)$$

Rectangular notation

$$\mathbf{v} = (v_x, v_y)$$

$$\mathbf{A} + \mathbf{B} = (A_x + B_x, A_y + B_y)$$

Converting a vector r

$$r_x = r \cos \theta$$

$$r_y = r \sin \theta$$

$$r = \sqrt{r_x^2 + r_y^2}$$

$$\theta = \arctan (r_y / r_x)$$

Chapter 3 Problems

Conceptual Problems

C.1 List three quantities that are represented by vectors.

C.2 Compare these two vectors: (5, 185°) and the negative of (5, 5°). Are they the same vector? Why or why not?

○ Yes ○ No

C.3 An aircraft carrier sails northeast at a speed of 6.0 knots. Its velocity vector is $\mathbf{v}$. What direction and speed would a ship with velocity vector $-\mathbf{v}$ have?

__________ knots

i. Northwest
ii. Northeast
iii. Southwest
iv. Southeast

C.4 Can the same vector have different representations in polar notation that use different angles? Explain.

○ Yes ○ No

C.5 A Boston cab driver picks up a passenger at Fenway Park, drops her off at the Fleet Center. Represent this displacement with the vector **D**. What is the displacement vector from the Fleet Center to Fenway Park?

○ **D** ○ 2**D** ○ 0 ○ −**D**

C.6 Does the multiplication of a scalar and a vector display the commutative property? This property states that the order of multiplication does not matter. So for example, if the multiplication of a scalar *s* and a vector **r** is commutative, then *s***r** = **r***s* for all values of *s* and **r**.

○ Yes ○ No

C.7 A small crab fishing boat travels from Colon City on the Pacific Ocean, through the Panama Canal, to Panama City on the Atlantic Ocean. A large cruise ship travels from Colon City to Panama City by sailing all the way around the Cape Horn at the southern tip of South America. Do these two voyages have equal displacement vectors?

○ Yes ○ No

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. Assume that the simulation is reset before each part and give each answer in the form (*x*, *y*). (a) What is the displacement to Ed's Fuel Depot? (b) What is the displacement to Joe's Diner? (c) What is the displacement to Silver's Gym?

(a) (__________ km, __________ km)
(b) (__________ km, __________ km)
(c) (__________ km, __________ km)

Section 1 - Scalars

1.1 The Earth has a mass of 5.97×10^{24} kg. The Earth's Moon has a mass of 7.35×10^{22} kg. How many Moons would it take to have the same mass as the Earth?

____

1.2 The volume of the Earth's oceans is approximately 1.4×10^{18} m^3. The Earth's radius is 6.4×10^{6} m. What percentage of the Earth, by volume, is ocean?

__________ %

1.3 Density is calculated by dividing the mass of an object by its volume. The Sun has a mass of 1.99×10^{30} kg and a radius of 6.96×10^{8} m. What is the average density of the Sun?

__________ kg/m^3

Section 3 - Polar notation

3.1 The tugboat Lawowa is returning to port for the day. It has a speed of 7.00 knots. The heading to the Lawowa's port is 31.0° west of north. If due east is 0°, what is the tug's heading as a vector in polar notation?

(__________ knots, __________ °)

3.2 An analog clock has stopped. Its hands are stuck displaying the time of 10 o'clock. The hour hand is 5.0 centimeters long, and the minute hand is 11 centimeters long. Write the position vector of the tip of the hour hand, in polar notation. Consider 3 o'clock to be 0°, and assume that the center of the clock is the origin.

(_________ cm, _________ °)

3.3 What is the polar notation for a vector that points from the origin to the point (0, 3.00)?

(_________ , _________ °)

Section 4 - Vector components and rectangular notation

4.1 (a) A hotdog vendor named Sam is walking from the southwest corner of Central park to the Empire State Building. He starts at the intersection of 59th Street and Eighth Avenue, walks 3 blocks east to 59th Street and Fifth Avenue, and then 25 blocks south to the Empire State Building at the Corner of 34th Street and Fifth Avenue. Write his displacement vector in rectangular notation with units of "blocks." Orient the axes so positive *y* is to the north, and positive *x* is to the east. (b) A stockbroker named Andrea makes the same trip in a cab that gets lost, and detours 50 blocks south to Washington Square before reorienting and finally arriving at the Empire State Building. Is her displacement any different?

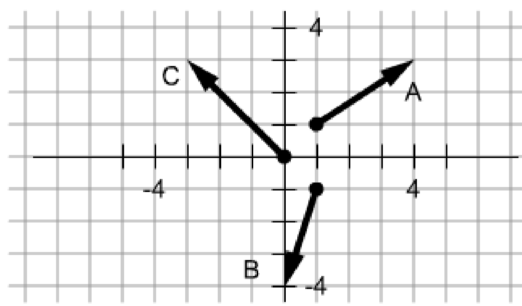
(a) (_________ , _________) blocks
(b) ○ Yes ○ No

4.2 A treasure map you uncovered while vacationing on the Spanish Coast reads as follows: "If me treasure ye wants, me hoard ye'll have, just follow thee directions these. Step to the south from Brisbain's Mouth, 5 paces through the trees. Then to the west, 10 paces ye'll quest, with mud as deep as yer knees. Then 3 paces more north, and dig straight down in the Earth, and me treasure, take it please." What is the displacement vector from Brisbain's Mouth to the spot on the Earth above the treasure? Consider east the positive *x* direction and north the positive *y* direction.

(_________ , _________) paces

4.3 Write the vectors labeled **A**, **B** and **C** with rectangular coordinates.

(a) A = (_________ , _________)
(b) B = (_________ , _________)
(c) C = (_________ , _________)



4.4 Consider these four vectors:
A goes from (0, 0) to (1, 2)
B from (1, −2) to (0, 2)
C from (−2, −1) to (−3, −3)
D from (−3, 1) to (−2, 3)
(a) Draw the vectors. Then answer the next two questions. (b) Which two vectors are equal? (c) Which vector is the negative of the two equal vectors?

(a) Submit anwer on paper.
(b) i. **A** and **B**
ii. **A** and **C**
iii. **A** and **D**
iv. **B** and **C**
v. **B** and **D**
vi. **C** and **D**
(c) i. **A**
ii. **B**
iii. **C**
iv. **D**

Section 5 - Adding and subtracting vectors graphically

5.1 Draw each of the following pairs of vectors on a coordinate system, using separate coordinate systems for parts a and b of the question. Then, on each coordinate system, also draw the vectors $-\mathbf{B}$, $\mathbf{A} + \mathbf{B}$, and $\mathbf{A} - \mathbf{B}$. Label all your vectors.
(a) $\mathbf{A} = (0, 5)$; $\mathbf{B} = (3, 0)$
(b) $\mathbf{A} = (4, 1)$; $\mathbf{B} = (2, -3)$

5.2 The Moon's orbit around the Earth is nearly circular, with an average radius of 3.8×10^5 kilometers from the Earth's center. It takes about 28 days for the Moon to complete one revolution around the Earth. (a) Put Earth at the origin of a coordinate system, and draw labeled vectors to represent the moon's position at 0, 7, 14, 21 and 28 days. (b) On a separate coordinate system, draw four labeled vectors representing the Moon's displacement from 0 days to 7 days, 7 days to 14 days, 14 days to 21 days, and 21 days to 28 days. (c) What is the sum of the four displacement vectors you drew in part "b"?

(a) Submit answer on paper.
(b) Submit answer on paper.
(c) __________

5.3 Consider the following vectors:
A goes from (0, 2) to (4, 2)
B from (1, −2) to (2, 1)
C from (−1, 0) to (0, 0)
D from (−3, −5) to (2, −2)
E from (−3, −2) to (−4, −5)
F from (−1, 3) to (−3, 0)
(a) Draw the vectors. Using your sketch and your knowledge of graphical vector addition and subtraction, which vector listed above is equal to: (b) **A** + **B**? (c) **E** − **F**? (d) −(**B** + **C**)? (e) Which two vectors sum to zero?

(a) Submit answer on paper.
(b)
i. A
ii. B
iii. C
iv. D
v. E
vi. F

(c)
i. A
ii. B
iii. C
iv. D
v. E
vi. F

(d)
i. A
ii. B
iii. C
iv. D
v. E
vi. F

(e)
i. A and D
ii. B and C
iii. B and E
iv. B and F
v. C and F
vi. E and F

5.4 The racing yacht *America* (USA) defeated the *Aurora* (England) in 1851 to win the 100 Guinea Cup. From the starting buoy the *America's* skipper sailed 400 meters at an angle 45° west of north, then 250 meters at an angle 30° east of north, and finally 350 meters at an angle 60° west of north. Draw the path of the *America* on a coordinate system as a set of vectors placed tip to tail. Then draw the total displacement vector.

Section 6 - Adding and subtracting vectors by components

6.1 Add the following vectors:
(a) (12, 5) + (6, 3)
(b) (−3, 8) + (6, −2)
(c) (3, 8, −7) + (7, 2, 17)
(d) (*a*, *b*, *c*) + (*d*, *e*, *f*)

(a) (__________ , __________)
(b) (__________ , __________)
(c) (__________ , __________ , __________)
(d)

6.2 Solve for the unknown variables:
(a) (a, b) + (3, 3) = (6, 7)
(b) (11, c) + (d, 2) = (−12, −3)
(c) (5, 5) + (e, 4) = (2, f)
(d) (4, −3) − (5, g) = (h, 2)

(a) a = __________ ; b = __________
(b) c = __________ ; d = __________
(c) e = __________ ; f = __________
(d) g = __________ ; h = __________

6.3 Solve for the unknown variables: (8, 3) + (b, 2) = (4, a).

(a) a = __________
(b) b = __________

6.4 Physicists model a magnetic field by assigning to every point in space a vector that represents the strength and direction of the field at that point. Two magnetic fields that exist in the same region of space may be added as vectors at each point to find the representation of their combined magnetic field. The "tesla" is the unit of magnetic field strength. At a certain point, magnet 1 contributes a field of (−6.4, 6.1, −3.7) tesla and magnet 2 contributes a field of (−1.1, −4.5, 8.6) tesla. What is the combined magnetic field at this point?

(__________ , __________ , __________) tesla

Section 9 - Multiplying rectangular vectors by a scalar

9.1 Perform the following calculations.
(a) 6(3, −1, 8)
(b) −3(−3, 4, −5)
(c) −a(a, b, c)
(d) −2(a, 5, c) + 6 (3, −b, 2)

(a) (__________ , __________ , __________)
(b) (__________ , __________ , __________)
(c)
(d)

9.2 Three vectors that are neither parallel nor antiparallel can be arranged to form a triangle if they sum to (0, 0). (a) What vector forms a triangle with (0, 3) and (3, 0)? (b) If you multiply all three vectors by the scalar 2, do they still form a triangle? (c) What if you multiply them by the scalar a?

(a) (__________ , __________)
(b) ○ Yes ○ No
(c) ○ Yes ○ No

Section 10 - Multiplying polar vectors by a scalar

10.1
Perform the following computations. Express each vector in polar notation with a positive magnitude and an angle between 0° and 360°.
(a) 2(4, 230°)
(b) −3(7, 20°)
(c) −4(8, 260°)

(a) (__________ , __________ °)
(b) (__________ , __________ °)
(c) (__________ , __________ °)

10.2 A chimney sweep is climbing a long ladder that leans against the side of a house. If the displacement of her feet from the base of the ladder is given by (2.1 ft, 65°) when she is on the third rung, what is the displacement of her feet from the base when she has climbed twice as far?

(__________ ft, __________ °)

Section 11 - Converting vectors from polar to rectangular notation

11.1 Express the following vectors in rectangular notation.
(a) **A** = (3.00, 25.0°)
(b) **B** = (17.0, 135°)
(c) **C** = (4.00, −185°)
(d) **D** = −(4.00, 185°)

(a) A = (________ , ________)
(b) B = (________ , ________)
(c) C = (________ , ________)
(d) D = (________ , ________)

11.2 Consider the polar vectors **A** = (5, 12°) and **B** = (65, 90°). Which points farther in the *x* direction?

○ **A** ○ **B**

11.3 In the *xy* plane, vector **A** is 6.40 cm long, and at an angle of 51.0° to the *x* axis. Write the vector in rectangular coordinates.

(________ , ________) cm

11.4 The vector **A**, in polar notation, is (45.0, 250°), and **B** is (70.0, 20.0°). What is **A** + **B** in rectangular coordinates?

(________ , ________)

Section 12 - Converting vectors from rectangular to polar notation

12.1 Express the following vectors in polar (r, θ) notation.
(a) **T** = (4.00, 4.00)
(b) **U** = (0, 3.50)
(c) **V** = (−3.00, 5.00)
(d) **W** = (−7.00, −9.00)

(a) T = (________ , ________ °)
(b) U = (________ , ________ °)
(c) V = (________ , ________ °)
(d) W = (________ , ________ °)

12.2 The vector **F** is (12.0, −3.00) in rectangular notation. What is this same vector in polar notation? Express the angle as a positive number between 0° and 360°.

(________ , ________ °)

12.3 An adventurous aardvark arduously ambles from (3.0, 2.0) m to (−4.0, −2.0) m. State this displacement in polar coordinates.

(________ m, ________ °)

Section 16 - Unit vectors

16.1 A helicopter takes off from the origin with a constant velocity of (16.0**i** + 18.0**j** + 4.00**k**) m/s. What is its position when it reaches an altitude of 1500 meters? Altitude is measured in the **k** direction.

(________ i + ________ j + ________ k) m

16.2 A dune buggy is driving across the desert with constant velocity, starting at an oasis. After 1.25 hours its displacement from the oasis is given by (60.0, 32.0) km. What is its displacement 5.00 hours later?

(________ , ________) km

16.3 Suppose that **u** = 4**i** − **j** + 6**k** and **v** = 4**i** + 3**j** + 8**k**. Find a vector **w**, so that **u** − 2**w** = **v**.

________ i + ________ j + ________ k

16.4 A vector lies in the *xz* plane, forms an angle of 30.0° with the *x* axis, and has a magnitude of 10.0. Its *z* component is positive. Write this vector in unit vector notation.

________ i + ________ k

16.5 Vector **v** is (12.0, −3.00, −4.00). Find a vector that points in the same direction as **v**, but has a magnitude of 1.

(________ , ________ , ________)

16.6 Find a vector with a magnitude of 1.00 that bisects the angle between the vectors 5.00**i** + 11.0**j** and 2.00**i** − 1.00**j**. Give your answer in rectangular coordinates.

(________ , ________)

Section 17 - Interactive summary problem: back to base

17.1 Use the information given in the interactive problem in this section to calculate the following values. (a) The displacement vector for the red ship in rectangular notation. (b) The displacement vector for the yellow ship in polar notation. (c) The scalar multiple for the purple ship's displacement vector. Test your answers using the simulation.

(a) (________ km, ________ km)
(b) (________ km, ________ °)
(c) ________ × (2.0 km, 1.0 km)

Additional Problems

A.1 Air traffic controllers use radar to keep track of the location of aircraft. The radar displays an aircraft's location in terms of the compass direction from the controller to the aircraft, and the horizontal distance along the ground between the two. In addition, the aircraft transmits its altitude to the controller automatically. Create a coordinate system by making east the positive *x* axis, north the positive *y* axis, and altitude the *z* axis. An aircraft is approaching at an altitude of 5.0 km, a horizontal distance of 35 km, and a bearing of 65° east of north. Write its position vector in rectangular coordinates.

(________ , ________ , ________) km

A.2 An air traffic controller has two aircraft on radar. The first is at an altitude of 0.500 km, a horizontal distance of 3.00 km measured along the ground to the point directly underneath the plane, and a bearing of 115° west of north. The second aircraft is at an altitude of 1.00 km, a horizontal distance of 8.00 km, and a bearing of 35.0° east of north. Write the displacement vector from aircraft 1 to aircraft 2 in "cylindrical" coordinates. That is, write a 3-component vector (r, θ, z) whose first two components are the polar coordinates of the horizontal displacement between the planes, and whose third coordinate is the vertical displacement between the planes. Consider east to be the positive *x* axis and north to be the positive *y* axis.

(________ km, ________ ° , ________ km)

A.3 A bird flies 5.00 m at 50.0° and then 3.00 m at −30.0°. What is the bird's total displacement in polar notation?

(________ m, ________ °)

A.4 Show by repeated addition that multiplying the vector **A** = (1, 2, 3) by 3 is the same as adding up 3 copies of the vector. What is the final vector?

(________ , ________ , ________)

chapter 4 Motion in Two and Three Dimensions

4.0 - Introduction

Imagine that you are standing on the 86th floor observatory of the Empire State Building, holding a baseball. A friend waits in the street below, ready to catch the ball. You toss it forward and watch it move in that direction at the same time as it plummets toward the ground. Although you have not thrown the ball downward at all, common sense tells you that your friend had better be wearing a well-padded glove!

When you tossed the ball, you subconsciously split its movement into two dimensions. You supplied the initial forward velocity that caused the ball to move out toward the street. You did not have to supply any downward vertical velocity. The force of gravity did that for you, accelerating the ball toward the ground. If you had wanted to, you could have simply leaned over and dropped the ball off the roof, supplying no initial velocity at all and allowing gravity to take over.

To understand the baseball's motion, you need to analyze it in two dimensions. Physicists use x and y coordinates to discuss the horizontal and vertical motion of the ball. In the horizontal direction, along the x axis, you supply the initial forward velocity to the ball. In the vertical direction, along the y axis, gravity does the work. The ball's vertical velocity is completely independent of its horizontal velocity. In fact, the ball will land on the ground at the same time regardless of whether you drop it straight off the building or hurl it forward at a Randy Johnson-esque 98 miles per hour.

To get a feel for motion in two dimensions, run the simulations on the right. In the first simulation, you try to drive a race car around a circular track by controlling its x and y component velocities separately, using the arrow keys on your keyboard. The right arrow increases the x velocity and the left arrow decreases it. The up arrow key increases the y velocity, and yes, the down arrow decreases it. Your mission is to stay on the course and, if possible, complete a lap using these keys.

Your car will start moving when you press any of the arrow keys. On the gauges, you can observe the x and y velocities of your car, as well as its overall speed. Does changing the x velocity affect the y velocity, or vice-versa? How do the two velocities seem to relate to the overall speed?

There is also a clock, so you can see which among your friends gets the car around the track in the shortest time. There is no penalty for driving your car off the track, though striking a wall is not good for your insurance rates. Press RESET to start over. Happy motoring!

In the second simulation, you can experiment with motion in two dimensions by firing the cannon from the castle. The cannon fires the cannonball horizontally from the top of a tower. You change the horizontal velocity of the cannonball by dragging the head of the arrow. Try to hit the two haystacks on the plain to see who is hiding inside.

As the cannonball moves, look at the gauges in the control panel. One displays the horizontal velocity of the cannonball, its displacement per unit time along the x axis. The other gauge displays the cannonball's vertical velocity, its displacement per unit time along the y axis. As you use the simulation, consider these important questions: Does the cannonball's horizontal velocity change as it moves through the air? Does its vertical velocity change? The simulation pauses when a cannonball hits the ground, and the gauges display the values from an instant before that moment.

The simulation also contains a timer that starts when the ball is fired and stops when it hits a haystack or the ground. Note the values in the timer as you fire shots of varying horizontal velocity. Does the ball stay in the air longer if you increase the horizontal velocity, or does it stay in the air the same amount of time regardless of that velocity?

The answers to these questions are the keys to understanding what is called projectile motion, motion where the acceleration occurs due to gravity alone. This chapter will introduce you to motion in two and three dimensions; projectile motion is one example of this type of motion.

4.1 - Displacement in two dimensions

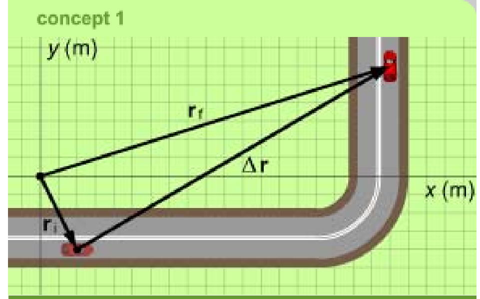
Straight-line motion means that an object is moving in only one dimension: north and south along Main Street, for example, or straight up and down, or along the x axis. Straight-line motion is the traditional starting point for the study of motion in physics.

The world is a little more complex than this. Objects do not move only along straight-line paths. A car moving around a curve, a baseball thrown from a tall building − these types of motion take place in more than one dimension.

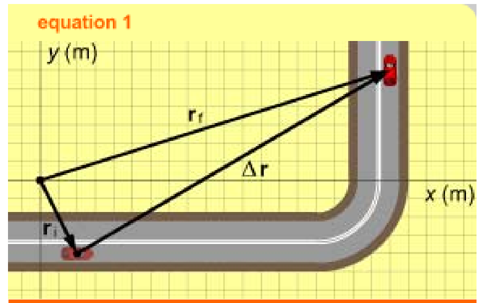
The car shown on the right moves in two dimensions simultaneously as it goes around the curve. A physicist (and you, shortly) would say that the car moves in both the x and y directions. Its position relative to the x and y axes changes as it moves.

The definition of displacement does not change when it is applied to motion in two or more dimensions. Displacement is always the net change in an object's position and it is always a vector. Since it is a vector, you need to know both its magnitude (the amount of the displacement) and its direction.

Position vectors can be used to represent an object's initial and final positions and they are particularly useful in analyzing multidimensional motion. At the right, the position vector $\mathbf{r}_i$ indicates the initial position of the car, and the position vector $\mathbf{r}_f$ points to the final position. The object's displacement equals the difference between those two positions. The vector $\Delta\mathbf{r}$ represents the displacement. The displacement is calculated by subtracting the initial position vector from the final position vector. This is stated as an equation in Equation 1 and used to solve the example problem on the right.



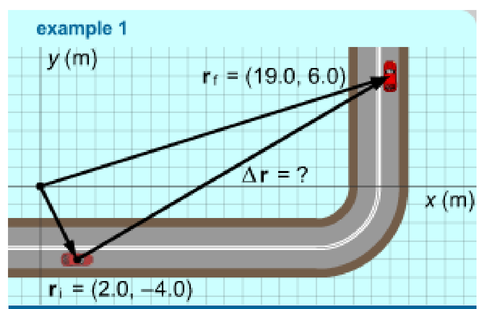
Displacement in two dimensions
Position vector r_i indicates initial location
Position vector r_f indicates final location
Displacement vector Δr points from initial to final location



Displacement in two dimensions

$$\Delta\mathbf{r} = \mathbf{r}_f - \mathbf{r}_i$$

$\Delta\mathbf{r}$ = displacement
$\mathbf{r}_f$ = final position
$\mathbf{r}_i$ = initial position



What is the displacement vector of the car?

$\Delta\mathbf{r} = \mathbf{r}_f - \mathbf{r}_i$

$\Delta\mathbf{r} = (19.0, 6.0)\text{ m} - (2.0, -4.0)\text{ m}$

$\Delta\mathbf{r} = (17.0, 10.0)\text{ m}$

4.2 - Velocity in two dimensions

Velocity is a vector quantity, meaning it contains two pieces of information: how fast something is traveling and in which direction. Both are crucial for understanding motion in multiple dimensions.

Consider the car on the track to the right. It starts out traveling parallel to the x axis at a constant speed. It then reaches a curve and continues to travel at a constant speed through the curve. Although its speed stays the same, its direction changes. Since velocity is defined by speed **and** direction, the change in direction means the car's velocity changes.

Using vectors to describe the car's velocity helps to illustrate its change in velocity. The velocity vector points in the direction of the car's motion at any moment in time. Initially, the car moves horizontally, and its velocity vector points to the right, parallel to the x axis. As the car goes around the curve, the velocity vector starts to point upward as well as to the right. You see this shown in the illustration for Concept 1.

When the car exits the curve, its velocity vector will be straight up, parallel to the y axis. Because the car is moving at a constant speed, the length of the vector stays the same: The speed, or magnitude of the vector, remains constant. However, the direction of the vector changes as the car moves around the curve.

Like any vector, the velocity vector can be written as the sum of its components, the velocities along the x and y axes. This is also shown in the Concept 1 illustration. The gauges display the x and y velocities. If you click on Concept 1 to see the animated version of the illustration, you will see the gauges constantly changing as the car rounds the bend. At the moment shown in the illustration, the car is moving at 17 m/s in the horizontal direction and 10 m/s in the vertical direction. The components of the vector shown also reflect these values. The horizontal component is longer than the vertical one.

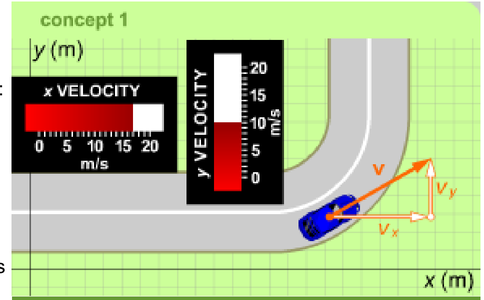
Equations 1 and 2 show equations useful for analyzing the car's velocity. Equation 1 shows how to break the car's overall velocity into its components. (These equations employ the same technique used to break any vector into its components.) The illustration shows the car's velocity vector. The angle θ is the angle the velocity vector makes with the positive x axis. The product of the cosine of that angle and the magnitude of the car's velocity (its speed) equals the car's horizontal velocity component. The sine of the angle times the speed equals the vertical velocity component.

The first equation in Equation 2 shows how to calculate the car's average velocity when its displacement and the elapsed time are known. The displacement $\Delta\mathbf{r}$ divided by the elapsed time Δt equals the average velocity.

The equations for determining the average velocity components when the components of the displacement are known are also shown in Equation 2. Dividing the displacement along the x axis by the elapsed time yields the horizontal component of the car's average velocity. The displacement along the y axis divided by the elapsed time equals the vertical component of the average velocity. A demonstration of these calculations is shown in Example 1.

The distinction between average and instantaneous velocity parallels the discussion of these two topics in the study of motion in one dimension. To determine the instantaneous velocity, $\Delta\mathbf{r}$ is measured during a very short increment of time and divided by that increment.

As with linear motion, the velocity vector points in the direction of motion. On the curved part of the track, the instantaneous velocity vector is tangent to the curve, since that is the direction of the car's motion at any instant in time. The average velocity vector points in the same direction as the displacement vector used to determine its value.

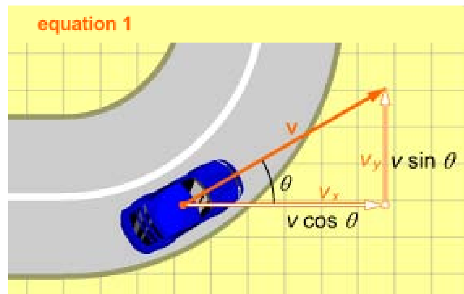


Velocity in two dimensions

Velocity has x and y components

Analyze x and y components separately

Two component vectors sum to equal total velocity



Components of velocity

$$v_x = v\cos\theta$$

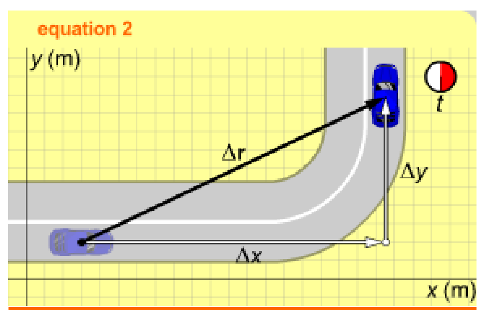
$$v_y = v\sin\theta$$

v = speed

θ = angle with positive x axis

v_x = x component of velocity

v_y = y component of velocity



Velocity from position, time

$$\bar{\mathbf{v}} = \frac{\Delta\mathbf{r}}{\Delta t}$$

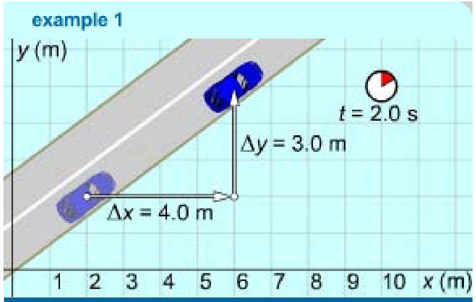
$$\bar{v}_x = \frac{\Delta x}{\Delta t}, \quad \bar{v}_y = \frac{\Delta y}{\Delta t}$$

$\mathbf{v}$ = velocity, $\mathbf{r}$ = position vector

Δt = elapsed time

Δx, Δy = x and y displacements

$\mathbf{v}$ becomes instantaneous as $\Delta t \to 0$



What are the *x* and *y* components of the car's average velocity?

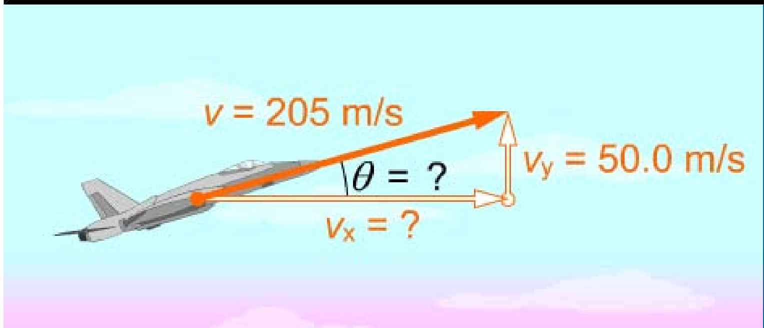
$$\bar{v}_x = \frac{\Delta x}{\Delta t}$$

$$\bar{v}_x = 4.0 \text{ m}/2.0 \text{ s} = 2.0 \text{ m/s}$$

$$\bar{v}_y = \frac{\Delta y}{\Delta t}$$

$$\bar{v}_y = 3.0 \text{ m}/2.0 \text{ s} = 1.5 \text{ m/s}$$

4.3 - Sample problem: velocity in two dimensions



A plane is flying at a speed of 205 m/s. The vertical component of the plane's velocity is 50.0 m/s.

What is the plane's horizontal velocity, and at what angle is the plane climbing?

Variables

speed	$v = 205$ m/s
vertical velocity	$v_y = 50.0$ m/s
horizontal velocity	v_x
angle	θ

What is the strategy?

1. Use trigonometry to solve for both the angle and the horizontal component of the velocity.

Physics principles and equations

$$v_y = v \sin\theta$$

$$v_x = v \cos\theta$$

Step-by-step solution

With the information given, we could solve first for either the plane's horizontal velocity or the angle relative to the horizontal. We choose to solve for the angle first.

Step	Reason
1. $v_y = v \sin\theta$	definition of vertical velocity component
2. $\sin\theta = \dfrac{v_y}{v}$	rearrange
3. $\sin\theta = \dfrac{50.0 \text{ m/s}}{205 \text{ m/s}} = 0.244$	enter values
4. $\theta = \arcsin(0.25) = 14.1°$	take arcsine to solve for θ

Now that we know the angle θ, we can use the other component equation to solve for the horizontal component of the plane's velocity.

Step	Reason
5. $v_x = v \cos\theta$	definition of horizontal velocity component
6. $v_x = (205 \text{ m/s})(\cos 14.1°)$	enter values
7. $v_x = 199$ m/s	evaluate

Note that we could also have used the Pythagorean theorem to solve for the horizontal velocity component without first solving for θ.

4.4 - Acceleration in two dimensions

Analyzing acceleration in two dimensions is analogous to analyzing velocity in two dimensions. Velocity can change independently in the horizontal and vertical dimensions. Because acceleration is the change in velocity per unit time, it follows that acceleration also can change independently in each dimension.

The cannonball shown to the right is fired with a horizontal velocity that remains constant throughout its flight. Constant velocity means zero acceleration. The cannonball has zero horizontal acceleration.

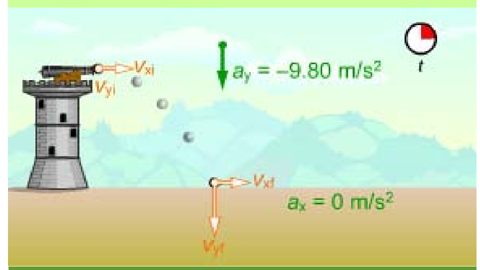
The cannonball starts with zero vertical velocity. Gravity causes its vertical velocity to become an increasingly negative number as the cannonball accelerates toward the ground. The vertical acceleration component due to gravity equals −9.80 m/s^2.

As with velocity, there are several ways to calculate the acceleration and its components.

The average acceleration can be calculated using the definition of acceleration, dividing the change in velocity by the elapsed time. The components of the average acceleration can be calculated by dividing the changes in the velocity components by the elapsed time. These equations are shown in Equation 1. Instantaneous acceleration is defined using the limit as Δt gets close to zero.

If the overall acceleration is known, in both magnitude and direction, you can calculate its x and y components by using the cosine and sine of the angle θ that indicates its direction. These equations are shown in Equation 2.

concept 1



Acceleration in two dimensions
Velocity can vary independently in *x*, *y* dimensions
Change in velocity = acceleration
Acceleration can also vary independently

equation 1

Acceleration in two dimensions

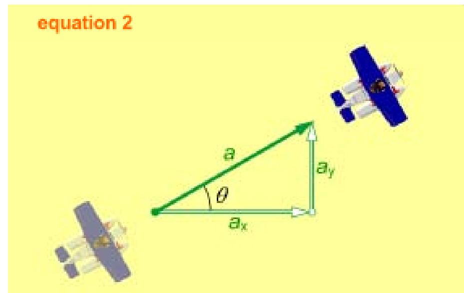
$$\bar{\mathbf{a}} = \frac{\Delta \mathbf{v}}{\Delta t}$$

$$\bar{a}_x = \frac{\Delta v_x}{\Delta t}, \quad \bar{a}_y = \frac{\Delta v_y}{\Delta t}$$

a = acceleration, **v** = velocity

Δt = elapsed time

Δv_x, Δv_y = velocity components



Acceleration in two dimensions

$$a_x = a \cos \theta$$

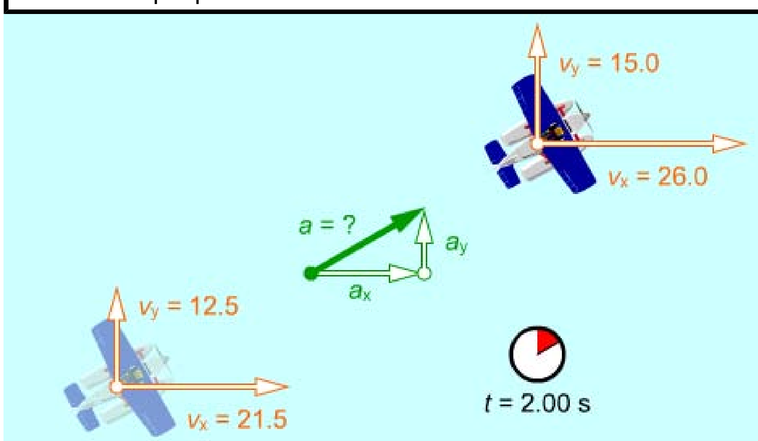
$$a_y = a \sin \theta$$

θ = angle with positive x axis

a_x = x component of acceleration

a_y = y component of acceleration

4.5 - Sample problem: acceleration in two dimensions



What are the components of the plane's acceleration? The magnitude and direction of its overall acceleration? Assume all are constant.

Variables

initial vertical velocity	v_{iy} = 12.5 m/s
final vertical velocity	v_{fy} = 15.0 m/s
initial horizontal velocity	v_{ix} = 21.5 m/s
final horizontal velocity	v_{fx} = 26.0 m/s
elapsed time	t = 2.00 s

What is the strategy?

1. Use the change in the x and y velocity components and the elapsed time to calculate the x and y components of the average acceleration. (A constant acceleration, specified in the problem, is the same as the average acceleration.)
2. Use the x and y components of the average acceleration and trigonometry to solve for the magnitude and direction of the acceleration.

Physics principles and equations

We will use the equations for the components of acceleration, omitting the bar notation because the acceleration is constant.

$$a_y = \frac{\Delta v_y}{t}$$

$$a_x = \frac{\Delta v_x}{t}$$

Step-by-step solution

First, we find the x and y components of the plane's acceleration. We do this by analyzing the components of the plane's velocity, starting with the vertical dimension.

Step	Reason
1. $a_y = \frac{\Delta v_y}{\Delta t}$	definition of vertical acceleration
2. $a_y = \frac{(15.0 \text{ m/s} - 12.5 \text{ m/s})}{2.00 \text{ s}}$	enter values
3. $a_y = 1.25 \text{ m/s}^2$	evaluate

Second, we find the other acceleration component, by analyzing the plane's horizontal velocity.

Step	Reason
4. $a_x = \frac{\Delta v_x}{\Delta t}$	definition of horizontal acceleration
5. $a_x = \frac{(26.0 \text{ m/s} - 21.5 \text{ m/s})}{2.00 \text{ s}}$	enter values
6. $a_x = 2.25 \text{ m/s}^2$	evaluate

Finally, we use the Pythagorean theorem and take the arccosine of the ratio a_x/a to solve for the plane's acceleration and its direction.

Step	Reason
7. $a^2 = a_x^2 + a_y^2$	Pythagorean theorem
8. $a^2 = (2.25)^2 + (1.25)^2$	enter values
9. $a = \sqrt{6.63} = 2.57 \text{ m/s}^2$	evaluate
10. $\cos\theta = \frac{a_x}{a} = 0.875$	trigonometry
11. $\theta = \arccos(0.875) = 29.0°$	arccos gives angle

We found that the plane is accelerating at an angle of 29.0° with the horizontal. The plane is **not** flying in that direction. The velocity vector changes direction due to the acceleration. Its initial direction is at 30.2° and its final direction is at 30.0°.

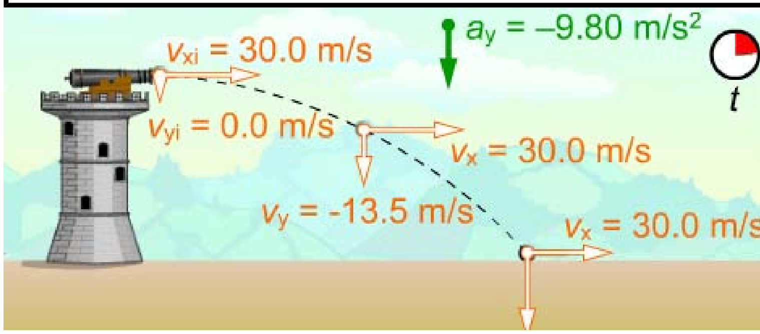
4.6 - Interactive checkpoint: acceleration in two dimensions

A runner is going around a track. She is initially moving with a velocity vector of (0.00, −8.00) m/s and her constant acceleration is (1.10, 1.10) m/s². What is her velocity 7.23 seconds later? Round the final velocity components to the nearest 0.01 m/s.

Answer:

$\mathbf{v}_f$ = (______ , ______) m/s

4.7 - Projectile motion



concept 1

Projectile velocity components
x and *y* velocity components
· *x* velocity constant (a_x = 0)
· *y* velocity changes (a_y= −9.80 m/s²)

Projectile motion: Movement determined by an object's initial velocity and the constant acceleration of gravity.

The path of a cannonball provides a classic example of projectile motion.

The cannonball leaves the cannon with an initial velocity and, ignoring air resistance, that initial velocity changes during the flight of the cannonball due solely to the acceleration due to the Earth's gravity.

The cannon shown in the illustrations fires the ball horizontally. After the initial blast, the cannon no longer exerts any force on the cannonball. The cannonball's horizontal velocity does not change until it hits the ground.

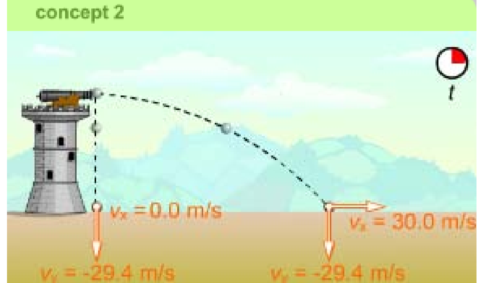
Once the cannonball begins its flight, the force of gravity accelerates it toward the ground. The force of gravity does not alter the cannonball's horizontal velocity; it only affects its vertical velocity, accelerating the cannonball toward the ground. Its y velocity has an increasingly negative value as it moves through the air.

The time it takes for the cannonball to hit the ground is completely unaffected by its horizontal velocity. It makes no difference if the cannonball flies out of the cannon with a horizontal velocity of 300 m/s or if it drops out of the cannon's mouth with a horizontal velocity of 0 m/s. In either case, the cannonball will take the same amount of time to land. Its vertical motion is determined solely by the acceleration due to gravity, −9.80 m/s².

The equations to the right illustrate how you can use the x and y components of the cannonball's velocity and acceleration to determine how long it will take to reach the ground (its *flight time*) and how far it will travel horizontally (its *range*). These are standard motion equations applied in the x and y directions. They hold true when the acceleration is constant, as is the case with projectile motion, where the acceleration along each dimension is constant.

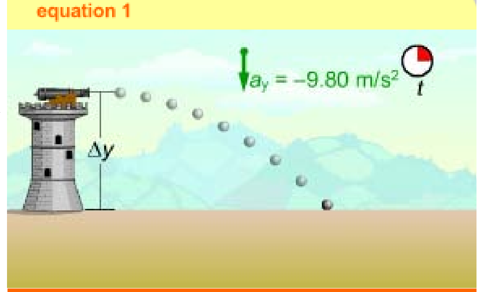
In Equation 1, we show how to determine how long it takes a projectile to reach the

concept 2



Projectile motion
Motion in one dimension independent of motion in other

equation 1



Projectile flight time

$v_{yi} = 0$

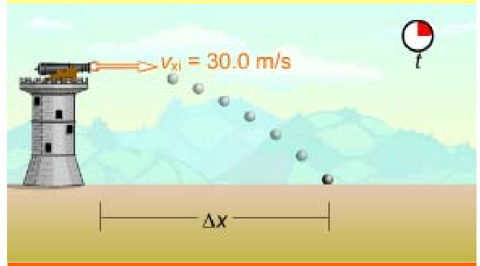
ground. This equation holds true when the initial vertical velocity is zero, as it is when a cannon fires horizontally. To derive the equation, we use a standard linear motion equation applied to the vertical, or y, dimension. You see that equation in the second line in Equation 1. We substitute zero for the initial vertical velocity and then solve for t, the elapsed time.

Equation 2 shows how to solve for the range. The horizontal displacement of the ball equals the product of its horizontal velocity and the elapsed time. Since its horizontal velocity does not change, it equals the initial horizontal velocity. Together, the two equations describe the flight of a projectile that is fired horizontally.

$$\Delta y = v_{yi}t+\frac{1}{2}a_y t^2$$
$$t = \sqrt{\frac{2\Delta y}{a_y}}$$

v_{yi} = initial y velocity
Δy = vertical displacement
$a_y = -9.80\ \text{m/s}^2$
t = time when projectile hits ground

equation 2

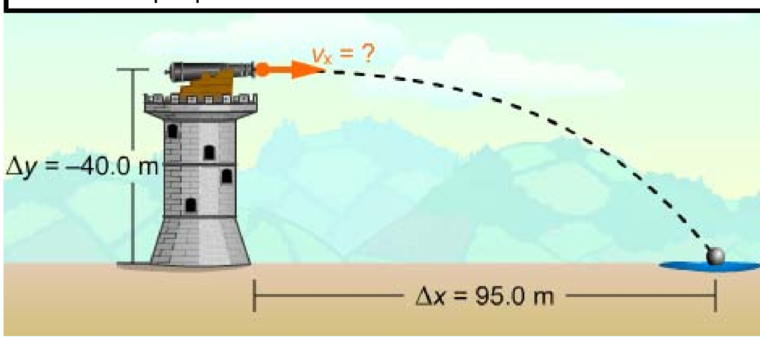


Projectile range

$$\Delta x = v_x t$$

Δx = horizontal displacement
v_x = horizontal velocity

4.8 - Sample problem: a horizontal cannon



What horizontal firing (muzzle) velocity splashes the cannonball into the pond?

Variables

vertical displacement	$\Delta y = -40.0$ m
vertical acceleration	$a_y = -9.80\ \text{m/s}^2$
horizontal displacement	$\Delta x = 95.0$ m
elapsed time	t
horizontal velocity	v_x

What is the strategy?

1. Calculate how long the cannonball remains in the air. This can be done with a standard motion equation.
2. Use the elapsed time and the specified horizontal displacement to calculate the required horizontal firing velocity.

Physics principles and equations

$$\Delta y = v_{yi}t+\frac{1}{2}a_y t^2$$
$$v_x = \frac{\Delta x}{t}$$

The amount of time the ball takes to fall is independent of its horizontal velocity.

Step-by-step solution

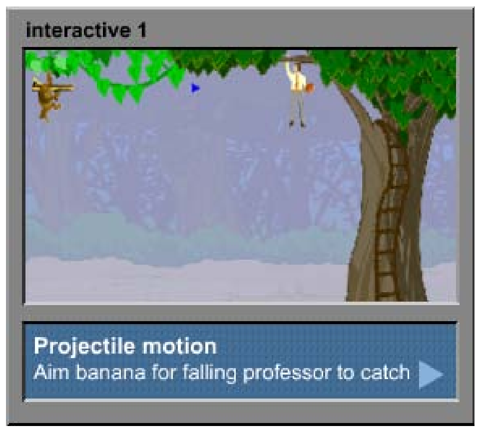
We start by determining how long the cannonball is in the air. We can use a linear motion equation to find the time it takes the cannonball to drop to the ground.

Step	Reason
1. $\Delta y = v_{yi}t + ½\, a_y t^2$	linear motion equation
2. $\Delta y = (0)t + ½\, a_y t^2$	initial vertical velocity zero
3. $t = \sqrt{\dfrac{2\Delta y}{a_y}}$	solve for time
4. $t = \sqrt{\dfrac{2(-40.0\text{ m})}{-9.80\text{ m/s}^2}}$	enter values
5. $t = 2.86$ s	evaluate

Now that we know the time the cannonball takes to fall to the ground, we can calculate the required horizontal velocity.

Step	Reason
6. $v_x = \dfrac{\Delta x}{t}$	definition of velocity
7. $v_x = \dfrac{95.0\text{ m}}{2.86\text{ s}}$	enter values
8. $v_x = 33.2$ m/s	divide

4.9 - Interactive problem: the monkey and the professor



The monkey at the right has a banana bazooka and plans to shoot a banana at a hungry physics professor. He has a glove to catch the banana. The professor is hanging from the tree, and the instant he sees the banana moving, he will drop from the tree in his eagerness to dine.

Can you correctly aim the monkey's bazooka so that the banana reaches the professor's glove as he falls? Should you aim the shot above, below or directly at the professor? As long as the banana is fired fast enough to reach the professor before it hits the ground, does its initial speed matter? (Note: The simulation has a minimum speed so the banana will reach the professor when correctly aimed.)

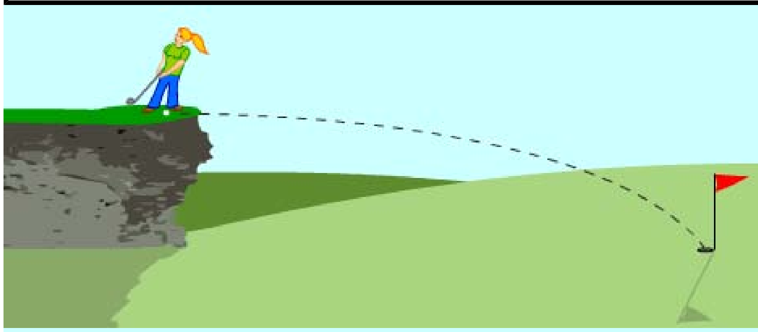
Give it a try in the interactive simulation to the right. No calculations are required to solve this problem. As you ponder your answer, consider the two key concepts of projectile motion: (1) all objects accelerate toward the ground at the same rate, and (2) the horizontal and vertical components of motion are independent. (The effect of air resistance is ignored.)

Aim the banana by dragging the vector arrow at the end of the bazooka. You can increase the firing speed of the banana by making the arrow longer, and you can change the angle at which the banana is fired by moving the arrow up or down. Stretching out the vector makes it easy to aim the banana.

To shoot the banana, press GO. Press RESET to try again. (Do not worry: We, too, value physics professors, so the professor will emerge unscathed.)

If you have trouble with this problem, review the section on projectile motion.

4.10 - Interactive checkpoint: golfing



A golfer is on the edge of a 12.5 m high bluff overlooking the eighteenth hole, which is located 67.1 m from the base of the bluff. She launches a horizontal shot that lands in the hole on the fly, and the gallery erupts into cheers. How long was the ball in the air? What was the ball's horizontal velocity? Take upward to be the positive *y* direction.

Answer:

t = ______ s

v_x = ______ m/s

4.11 - Projectile motion: juggling

Juggling is a form of projectile motion in which the projectiles have initial velocities in both the vertical and horizontal dimensions. This motion takes more work to analyze than when a projectile's initial vertical velocity is zero, as it was with the horizontally fired cannonball.

Jugglers throw balls from one hand to the other and then back. To juggle multiple balls, the juggler repeats the same simple toss over and over. An experienced juggler's ability to make this basic routine seem so effortless stems from the fact that the motion of each ball is identical. The balls always arrive at the same place for the catch, and in roughly the same amount of time.

A juggler throws each ball with an initial velocity that has both x and y components. Ignoring the effect of air resistance, the x component of the velocity remains constant as the ball moves in the air from one hand to another.

The initial y velocity is upward, which means it is a positive value. At all times, the ball accelerates downward at −9.80 m/s^2. This means the ball's velocity decreases as it rises until it has a vertical velocity of zero. The ball then accelerates back toward the ground. When the ball plops down into the other hand, the magnitude of the y velocity will be the same as when the ball was tossed up, but its sign will be reversed. If the ball is thrown up with an initial y velocity of +5 m/s, it will land with a y velocity of −5 m/s. This symmetry is due to the constant rate of vertical acceleration caused by gravity.

Because the vertical acceleration is constant, the ball takes as much time to reach its peak of motion as it takes to fall back to the other hand. This also means it has covered half of the horizontal trip when it is at its peak. The path traced by the ball is a *parabola*, a shape symmetrical around its midpoint.

We said the juggler chooses the *y* velocity so that the ball takes the right amount of time to rise and then fall back to the other hand. What, exactly, is the right amount of time? It is the time needed for the ball to move the horizontal distance between the juggler's hands.

If a juggler's hands are 0.5 meters apart, and the horizontal velocity of the ball is 0.5 m/s, it will take one second for the ball to move that horizontal distance. (Remember that the horizontal velocity is constant.) The juggler must throw the ball with enough vertical velocity to keep it in the air for one second.

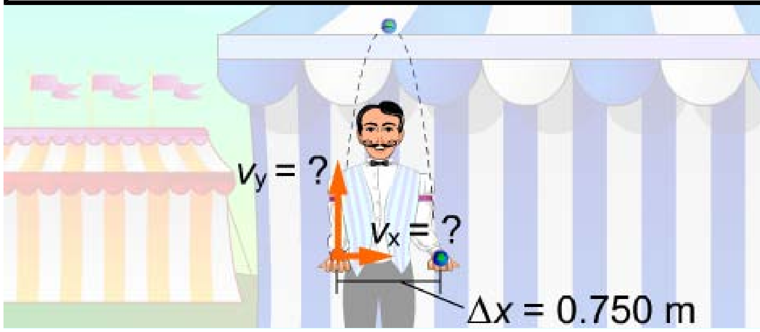
concept 1

Projectile motion: *y* velocity
y velocity = zero at peak
Initial *y* velocity equal but opposite to final *y* velocity

concept 2

Projectile motion: *x* velocity
x velocity constant

4.12 - Sample problem: calculating initial velocity in projectile motion



A juggler throws each ball so it hangs in the air for 1.20 seconds before landing in the other hand, 0.750 meters away.

What are the initial vertical and horizontal velocity components?

Variables

elapsed time	$t = 1.20$ s
horizontal displacement	$\Delta x = 0.750$ m
initial vertical velocity	v_y
horizontal velocity	v_x
acceleration due to gravity	$a_y = -9.80$ m/s^2

What is the strategy?

1. Use a linear motion equation to determine the initial vertical velocity that will result in a final vertical velocity of 0 m/s after 0.60 s (half the total time of 1.20 s).
2. Use the definition of velocity to calculate the horizontal velocity, since the displacement and the elapsed time are both provided.

Physics principles and equations

We rely on two concepts. First, the motion of a projectile is symmetrical, so half the time elapses on the way up and the other half on the way down. Second, the vertical velocity of a projectile is zero at its peak.

We also use the two equations listed below.

$$v_{yf} = v_{yi} + a_y t$$

$$v_x = \frac{\Delta x}{t}$$

Step-by-step solution

We start by calculating the vertical component of the initial velocity. We use the fact that gravity will slow the ball to a vertical velocity of 0 m/s at the peak, 0.60 seconds (halfway) into the flight.

Step	Reason
1. $v_{yf} = v_{yi} + a_y t$	motion equation
2. $0 = v_{yi} + a_y t$	vertical velocity at peak equals zero
3. $v_{yi} = -a_y t$	rearrange
4. $v_{yi} = -(-9.80 \text{ m/s}^2)(0.600 \text{ s})$	enter values
5. $v_{yi} = 5.89$ m/s	multiply

Now we solve for the horizontal velocity component. We could have solved for this first; these steps require no results from the steps above.

Step	Reason
6. $v_x = \frac{\Delta x}{t}$	definition of velocity
7. $v_x = \frac{0.750 \text{ m}}{1.20 \text{ s}}$	enter values
8. $v_x = 0.625$ m/s	divide

From our perspective, the horizontal velocity of the ball when it is going from our left to our right is positive. From the juggler's, it is negative.

4.13 - Interactive problem: the monkey and the professor, part II

The monkey is at it again with his banana bazooka. Another hungry professor drops from the tree the instant the banana leaves the tip of the bazooka. (Are professors paid enough? Answer: No.)

Can you correctly aim the banana so that it hits the professor's glove as she falls? Should you aim the shot above, below or directly at the glove? Does the banana's initial speed matter? (Note: The simulation has a minimum speed so the banana will reach the professor when correctly aimed.) Air resistance is ignored in this simulation.

Give it a try in the interactive simulation to the right. No calculations are needed to solve this problem.

To find an answer, think about the key concepts of projectile motion: Objects accelerate toward the ground at the same rate, and the horizontal and vertical components of motion are independent.

Once you launch the simulation, aim the banana by dragging the vector arrow at the end of the bazooka. You can increase the firing speed of the banana by making the arrow longer. Moving the arrow up or down changes the angle at which the banana is fired. Stretching out the vector makes it easy to aim the banana.

To shoot the banana, press GO. Press RESET to try again. Keep trying until the banana reaches the professor!

4.14 - Projectile motion: aiming a cannon

Aiming a projectile, step 1
Start with initial angle, speed
Separate into *x* and *y* components

Aiming a projectile, step 2
Use initial *y* velocity and a_y to calculate flight time
Use initial *x* velocity and flight time to calculate range

The complexities of correctly aiming artillery pieces have challenged leaders as famed as Napoleon and President Harry S. Truman. Because a cannon typically fires projectiles at a particular speed, aiming the cannon to hit a target downfield involves adjusting the cannon's angle relative to the ground. If you break the motion into components, you can determine how far a projectile with a given speed and angle will travel.

To determine when and where a cannonball will land, you must consider horizontal and vertical motion separately. To start, convert its initial speed and angle into x and y velocity components. The horizontal velocity will equal the initial speed of the ball multiplied by the cosine of the angle at which the cannonball is launched. The horizontal velocity will not change as the cannonball flies toward the target.

The initial y velocity equals the initial speed times the sine of the launch angle. The y velocity is not constant. It changes at the rate of −9.80 m/s^2. When the cannonball lands at the same height at which it was fired, its final y velocity is equal but opposite to its initial y velocity.

The initial y velocity of the projectile determines how long it stays in the air. As mentioned, the cannonball lands with a final y velocity equal to the negative of the initial y velocity, that is, $v_{yf} = -v_{yi}$. This means the **change** in y velocity equals $-2v_{yi}$. Knowing this, and the value for acceleration due to gravity, enables us to rearrange a standard motion equation ($v_f = v_i + at$) and solve for the elapsed time. The equation for the flight time of a projectile is the third one in Equation 1.

Once you know how long the ball stays in the air, you can determine how far it travels by multiplying the horizontal velocity by the ball's flight time. This is the final equation on the right.

You can use these equations to solve projectile motion problems, but understanding the analysis that led to the equations is more important than knowing the equations. Recall the basic principles of projectile motion: The x velocity is constant, the y velocity changes at the rate of −9.80 m/s^2, and the projectile's final y velocity is the opposite of its initial y velocity.

equation 1

Projectile equations

$v_x = v \cos\theta$

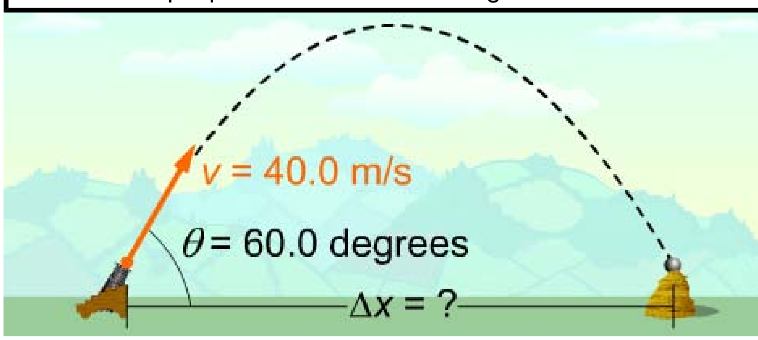
$v_y = v \sin\theta$

$t = -2v_y / a_y$ (same-height landing)

$\Delta x = v_x t$

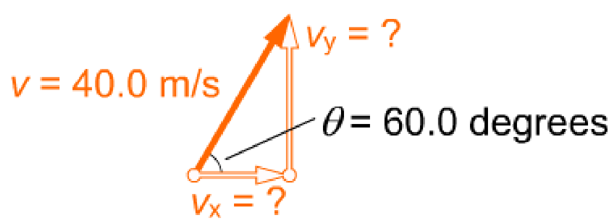
v_x = x velocity
v_y = initial y velocity
t = time projectile is in air
Δx = horizontal displacement
$a_y = -9.80\ \text{m/s}^2$

4.15 - Sample problem: a cannon's range



How far away is the haystack from the cannon?

Draw a diagram



Variables

speed	$v = 40.0$ m/s
angle	$\theta = 60.0°$
initial y velocity	v_{yi}
final y velocity	v_{yf}
x velocity	v_x
elapsed time	t
horizontal displacement	Δx
acceleration due to gravity	$a_y = -9.80\ \text{m/s}^2$

What is the strategy?

1. Use trigonometry to determine the x and y components of the cannonball's initial velocity.
2. The final y velocity is the opposite of the initial y velocity. Use that fact and a linear motion equation to determine how long the cannonball is in the air.
3. The x velocity is constant. Rearrange the definition of constant velocity to solve for horizontal displacement (range).

Physics principles and equations

The projectile's final y velocity is the opposite of its initial y velocity. The x velocity is constant.

We use the following motion equations.

$$v_{yf} = v_{yi} + a_y t$$

$$v_x = \frac{\Delta x}{t}$$

Step-by-step solution

Use trigonometry to determine the x and y components of the initial velocity from the initial speed and the angle.

Step	Reason
1. $v_y = v \sin\theta$	trigonometry
2. $v_y = (40.0\ \text{m/s}) \sin 60.0°$	enter values
3. $v_y = 34.6\ \text{m/s}$	evaluate
4. $v_x = v \cos\theta$	trigonometry
5. $v_x = (40.0\ \text{m/s}) \cos 60.0°$	enter values
6. $v_x = 20.0\ \text{m/s}$	evaluate

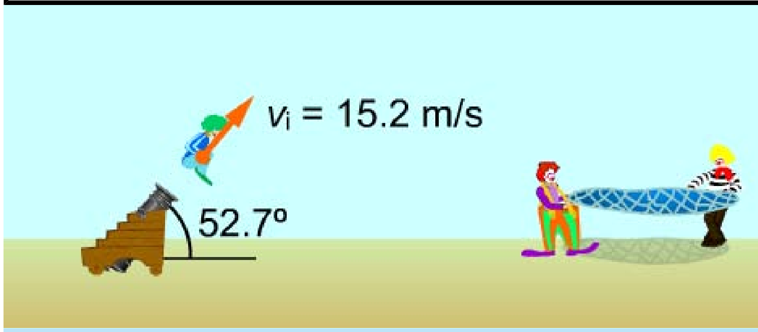
Now we focus on the vertical dimension of motion, using the initial y velocity to determine the time the cannonball is in the air. We calculated the initial y velocity in step 3.

Step	Reason
7. $v_{yf} = v_{yi} + a_y t$	linear motion equation
8. $-v_{yi} = v_{yi} + a_y t$	final y velocity is negative of initial y velocity
9. $-2v_{yi} = a_y t$	rearrange
10. $t = \dfrac{-2v_{yi}}{a_y}$	solve for time
11. $t = \dfrac{-2(34.6\ \text{m/s})}{-9.80\ \text{m/s}^2}$	enter values
12. $t = 7.07\ \text{s}$	evaluate

Now we use the time and the x velocity to solve for the cannonball's range. We calculated the constant x velocity in step 6.

Step	Reason
13. $v_x = \dfrac{\Delta x}{t}$	definition of velocity
14. $\Delta x = v_x t$	rearrange
15. $\Delta x = (20.0\ \text{m/s})(7.07\ \text{s})$	enter values
16. $\Delta x = 141\ \text{m}$	multiply

4.16 - Interactive checkpoint: clown cannon

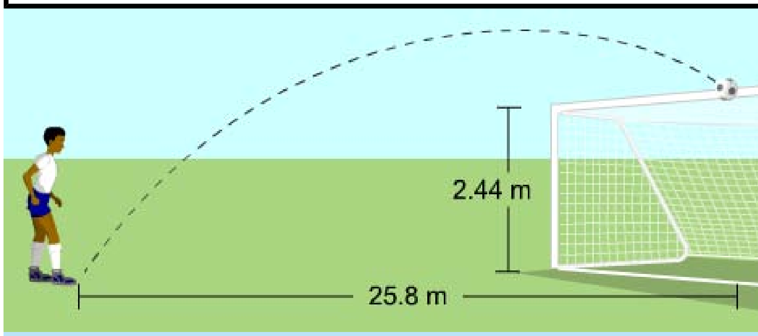


A clown in a circus is about to be shot out of a cannon with a muzzle velocity of 15.2 m/s, aimed at 52.7° above the horizontal. How far away should his fellow clowns position a net to ensure that he lands unscathed? The net is at the same height as the mouth of the cannon.

Answer:

Δx = [ ] m

4.17 - Interactive checkpoint: soccer kick



A professional soccer player is 25.8 m from the goal, and kicks a hard shot from ground level. The ball hits the crossbar on its way down, 2.44 m above the ground, 1.98 s after it was kicked. What were the *x* and *y* components of the ball's initial velocity? Take upward to be the positive *y* direction.

Answer:

v_x = [ ] m/s

v_{iy} = [ ] m/s

4.18 - Interactive problem: the human cannonball

The world record for the farthest flight by a human shot out of a cannon was set in Pennsylvania in 1998 by David "Cannonball" Smith, Sr., who traveled 56.54 meters across a lake, landing in a net.

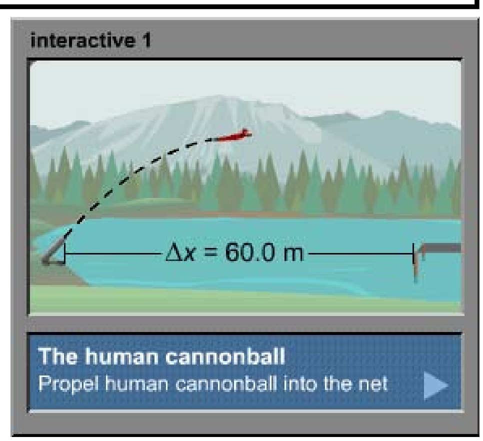
Here is your chance to break the world record. The net is 60.0 m away from the cannon's mouth (and at the same height). Can you calculate the correct firing speed and angle to propel the human cannonball into the net?

As with other projectile problems, you will need to figure out which combination of horizontal and initial vertical velocities will propel the human cannonball the correct horizontal distance in the time it takes the daredevil to rise and fall back to Earth.

HINT: There are an infinite number of answers to this problem. Try setting a reasonable value for the firing speed (try a few if necessary) and see if you can calculate the correct angle for that speed.

You will need to solve the problem using the components of the velocity, and then use trigonometry to determine the speed and angle required in this simulation. Enter your answer for speed to the nearest 0.1 m/s, and your answer for angle to the nearest 0.1 degree.

If you have trouble with this simulation, refer to the sections on projectile motion.



4.19 - Interactive problem: test your juggling!

Much of this chapter focuses on projectile motion: specifically, how objects move in two dimensions. If you have grasped all the concepts, you can use what you have learned to make the person at the right juggle.

The distance between the juggler's hands is 0.70 meters and the acceleration due to gravity is −9.80 m/s^2. You have to calculate the initial x

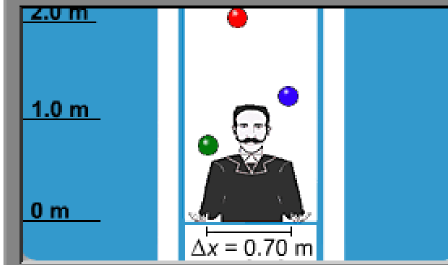
and y velocities to send each ball from one hand to the other. If you do so correctly, he will juggle three balls at once.

There are many possible answers to this problem. A good strategy is to pick an initial x or y component of the velocity, and then determine the other velocity component so that the balls, once thrown, will land in the juggler's opposite hand. You want to pick an initial y velocity above 2.0 m/s to give the juggler time to make his catch and throw. For similar reasons, you do not want to pick an initial x velocity that exceeds 2.0 m/s.

Make your calculations and then click on the diagram to the right to launch the simulation. Enter the values you have calculated to the nearest 0.1 m/s and press the GO button. Do not worry about the timing of the juggler's throws. They are calculated for you automatically.

If you have difficulty with this problem, refer to the sections on projectile motion.

interactive 1



The Juggler
Use physics to astound your friends

4.20 - The range and elevation equations

Often, with cannons and other devices that launch projectiles, the initial projectile speed is fixed and known. The angle above the horizontal is what dictates how far the projectile will travel when it lands at the same height at which it is fired.

This means the angle above the horizontal, the shooting angle or *elevation*, determines the range (the horizontal distance) traveled by the projectile.

In this section, we derive an equation for the range of the projectile, rearrange it to solve for the shooting angle, and then comment on a few of its implications. You see the range equation in Equation 1 and the elevation equation in Equation 2.

The effects of air resistance are ignored.

Variables

Since the projectile starts and ends at the same height, the vertical displacement is zero.

horizontal displacement	Δx
vertical displacement	$\Delta y = 0$ m
initial speed of projectile	v
initial velocity component in *x* direction	v_x
initial velocity component in *y* direction	v_y
flight time of projectile	t
acceleration due to gravity	$a_y = -9.80$ m/s^2
shooting angle (elevation)	θ

Strategy

1. Use the initial speed in the y direction and a motion equation to find the flight time of the projectile.
2. Then use the initial speed in the x direction, a motion equation and the flight time to find the range of the projectile.
3. With the help of trigonometry, express the range in terms of the angle. This is the range equation.
4. Finally, solve the range equation for the angle to get the elevation equation.

Physics principles and equations

An equation derived earlier to determine the flight time of a projectile

$$t = -2v_y/a_y$$

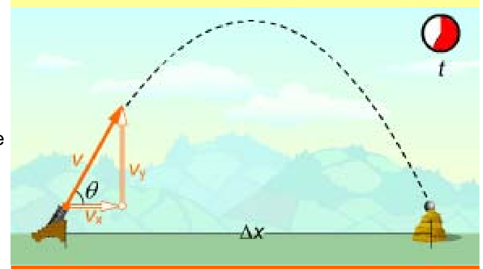
Since there is no horizontal acceleration

$$\Delta x = v_x t$$

The velocity components are calculated using these equations

$$v_x = v \cos \theta$$
$$v_y = v \sin \theta$$

equation 1

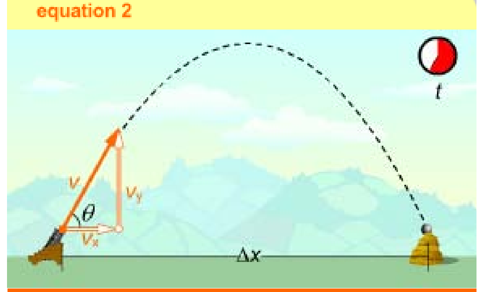


Derivation 1: the range equation

$$\Delta x = \frac{-v^2 \sin 2\theta}{a_y}$$

Δx = horizontal displacement (range)
v = initial velocity (firing speed)
θ = shooting angle (elevation)
a_y = acceleration due to gravity

equation 2



Derivation 2: the elevation equation

$$\theta = ½\arcsin\left(\frac{-a_y \Delta x}{v^2}\right)$$

Mathematics facts

We will use the trigonometric identity,

$$\sin 2\theta = 2\sin\theta\cos\theta$$

Step-by-step derivations

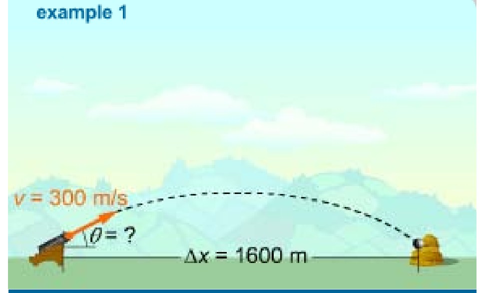
The range equation (derivation 1). We use the flight time equation stated above and the definition of velocity to determine the range as a function of the initial velocity components and the vertical acceleration.

Step	Reason
1. $\Delta x = v_x t$	definition of velocity
2. $\Delta x = v_x(-2v_y/a_y)$	substitute flight time equation into step 1
3. $\Delta x = \dfrac{-2v_y v_x}{a_y}$	rearrange

In the following steps we use trigonometric relationships to find the range of the projectile in terms of the initial speed and angle of elevation.

Step	Reason
4. $\Delta x = \dfrac{-2(v\sin\theta)(v\cos\theta)}{a_y}$	substitute given equations
5. $\Delta x = \dfrac{-v^2(2\sin\theta\cos\theta)}{a_y}$	rearrange
6. $\Delta x = \dfrac{-v^2\sin 2\theta}{a_y}$ (the range equation)	trigonometric identity

example 1



What is the correct elevation?

$$\theta = ½\arcsin\left(\frac{-a_y\Delta x}{v^2}\right)$$

$$= \frac{1}{2}\arcsin\left(\frac{-(-9.80\text{ m/s}^2)(1600\text{m})}{(300\text{ m/s})^2}\right)$$

$\theta = ½ \arcsin(0.174)$

$\theta = ½(10°) = 5°$ (forward shot)

$\theta = 90° - 5° = 85°$ (lob shot)

The elevation equation (derivation 2). Above, we derived the range equation. Now we solve this equation for the elevation or shooting angle, θ.

Step	Reason
7. $\sin 2\theta = \dfrac{-a_y\Delta x}{v^2}$	rearrange
8. $2\theta = \arcsin\left(\dfrac{-a_y\Delta x}{v^2}\right)$	arcsin of both sides
9. $\theta = ½\arcsin\left(\dfrac{-a_y\Delta x}{v^2}\right)$	solve for θ

For a given initial velocity the greatest range occurs when the factor $\sin 2\theta$ in the range equation is as large as possible. The sine function reaches a maximum value of 1 when its argument is $90°$. If 2θ is $90°$, then the angle θ that results in the maximum projectile range is $45°$. This fact has been taken advantage of since ancient times.

It is a fact of trigonometry that any two angles that add up to $180°$ have the same sine value. Therefore, shooting a projectile at $70°$ ($2\theta = 140°$) results in the same range as $20°$ ($2\theta = 40°$), because $\sin 2\theta$ is the same in both cases, and $\sin 2\theta$ is the factor that appears in the range equation. Another way of putting this is to say that a projectile fired at an angle θ will travel the same range as a projectile fired at $90° - \theta$. The high-angle shot is called a *lob* shot, while the low-angle shot is called a *forward* shot. In gunnery, mortars are generally used for lob shots, and "guns" are used for forward shots. Howitzers may be used for either.

As mentioned above, these equations ignore the effects of air resistance (sometimes called "drag"). The mathematics needed to take air resistance into account is complicated. It is easier to use a spreadsheet to break an object's motion down into small increments of time. You can see an example of such a spreadsheet by clicking here. In the spreadsheet, you can set the angle, initial height, and initial speed for two projectiles, one of which is subject to air resistance, and see a graph of their projectile paths. You can also control the "drag constant" that tells how much air resistance affects the motion. The shape and surface of an object determine its drag constant.

Reference frame: A coordinate system used to make observations.

The choice of a reference frame determines the perception of motion. A reference frame is a coordinate system used to make observations. If you stand next to a lab table and hold out a meter stick, you have established a reference frame for making observations.

The choice of reference frames was a minor issue when we considered juggling: We chose to measure the horizontal velocity of a ball as you saw it when you stood in front of the juggler. The coordinate system was established using your position and orientation, assuming you were stationary relative to the juggler.

As the juggler sees the horizontal velocity of the ball, however, it has the same magnitude you measure, but is opposite in sign. It does so because when you see it moving from your left to your right, he sees it moving from his right to his left. If you measure the velocity as 1.1 m/s, he measures it as −1.1 m/s.

In the analysis of motion, it is commonly assumed that you, the observer, are standing still. To pursue this further, we ask you to sit or stand still for a moment. Are you moving? Likely you will answer: “No, you just asked me to be still!”

That response is true for what you are implicitly using as your reference frame, your coordinate system for making measurements. You are implicitly using the Earth’s surface.

But from the perspective of someone watching from the Moon, you are moving due to the Earth’s rotation and orbital motion. Imagine that the person on the Moon wanted to launch a rocket to pick you up. Unless the person factored in your velocity as the Earth spins about its axis, as well as the fact that the Earth orbits the Sun and the Moon orbits the Earth, the rocket surely would miss its target. If you truly think you are stationary here on Earth, you must also conclude that the entire universe revolves around the Earth (a dubious conclusion, though a common one for centuries).

concept 1

Reference frames
System for observing motion

concept 2

Reference frames
Measurements of motion defined by reference frame

Reference frames define your perception and measurements of motion. If you are in a car moving at 80 km/h, another car moving alongside you with the same velocity will appear to you as if it is not moving at all. As you drive along, objects that you ordinarily think of as stationary, such as trees, seem to move rapidly past you. On the other hand, someone sitting in one of the trees would say the tree is stationary and you are the one moving by.

A reference frame is more than just a viewpoint: It is a coordinate system used to make measurements. For instance, you establish and use a reference frame when you do lab exercises. Consider making a series of measurements of how long it takes a ball to roll down a plane. You might say the ball’s starting point is the top of the ramp. Its x position there is 0.0 meters. You might define the surface of the table as having a y position of 0.0 meters, and the ball’s initial y position is its height above the table. Typically you consider the plane and table to be stationary, and the ball to be moving.

Two reference frames are shown on the right. One is defined by Joan, the woman standing at a train station. As the illustration in Concept 1 shows, from Joan’s perspective, she is stationary and the train is moving to the right at a constant velocity.

Another reference frame is defined by the perspective of Ted who is inside the train, and considers the train stationary. This reference frame is illustrated in Concept 2. Ted in the train perceives himself as stationary, and would see Joan moving backward at a constant velocity. He would assign Joan the velocity vector shown in the diagram.

It is important to note there is no correct reference frame; Joan cannot say her reference frame is better than the reference frame used by Ted. Measurements of velocity and other values made by either observer are equally valid.

Reference frames are often chosen for the sake of convenience (choosing the Earth’s surface, not the surface of Jupiter, is a logical choice for your lab exercises). Once you choose a reference frame, you must use it consistently, making all your measurements using that reference frame’s coordinate system. You cannot measure a ball’s initial position using the Earth’s surface as a reference frame, and its final position using the surface of Jupiter, and still easily apply the physics you are learning.

4.22 - Relative velocity

Observers in reference frames moving past one another may measure different velocities for the same object. This concept is called *relative velocity*.

In the illustrations to the right, two observers are measuring the velocity of a soccer ball, but from different vantage points: The man is standing on a moving train, while the woman is standing on the ground. The man and the woman will measure different velocities for the soccer ball.

Let’s discuss this scenario in more depth. Fred is standing on a train car and kicks a ball to the right. The train is moving along the track at a constant velocity. The train is Fred’s reference frame, and, to him, it is stationary. In Fred’s frame of reference, his kick causes the ball to move at a constant velocity of positive 10 m/s.

The train is passing Sarah, who is standing on the ground. Her reference frame is the ground. From her perspective, the train with the man on it moves by at a constant velocity of positive 5 m/s.

What velocity would Sarah measure for the soccer ball in her reference frame? She adds the velocity vector of the train, 5 m/s, to the velocity vector of the ball as measured on the train, 10 m/s. The sum is positive 15 m/s, pointing along the horizontal axis. Summing the velocities determines the velocity as measured by Sarah.

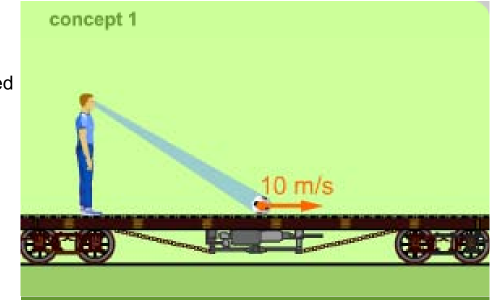
Note that there are two different answers for the velocity of one ball. Each answer is correct in the reference frame of that observer. For someone standing on the ground, the ball moves at 15 m/s, and for someone standing on the train, it moves at 10 m/s.

The equation in Equation 1 shows how to relate the velocity of an object in one reference frame to the velocity of an object in another frame. The variable $\mathbf{v}_{\mathrm{OA}}$ is the velocity vector of the object as measured in reference frame A (which in this diagram is the ground). The variable $\mathbf{v}_{\mathrm{OB}}$ is the velocity of an object as measured in reference frame B (which in this diagram is the train). Finally, the variable $\mathbf{v}_{\mathrm{BA}}$ is the velocity of frame B (the train) relative to frame A (the ground). $\mathbf{v}_{\mathrm{OA}}$ is the **vector** sum of $\mathbf{v}_{\mathrm{OB}}$ and $\mathbf{v}_{\mathrm{BA}}$.

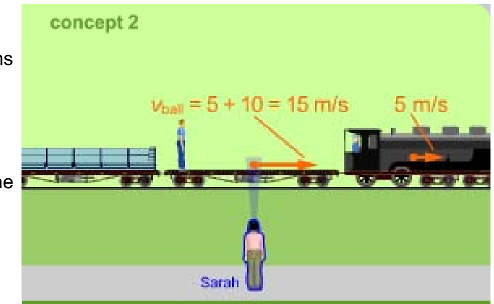
An important caveat is that this equation can be used to solve relative velocity problems only when the frames are moving at constant velocity relative to one another. If one or both frames are accelerating, the equation does not apply.

We use this equation in the example problem on the right. Now Sarah sees the train moving in the opposite direction, at **negative** 5 m/s. It is negative because to Sarah, the train is moving to the left along the *x* axis. Fred is on the train, again kicking the ball from left to right as before. Here he kicks the ball at +5 m/s, as measured in his reference frame. To Sarah, how fast and in what direction is the ball moving now?

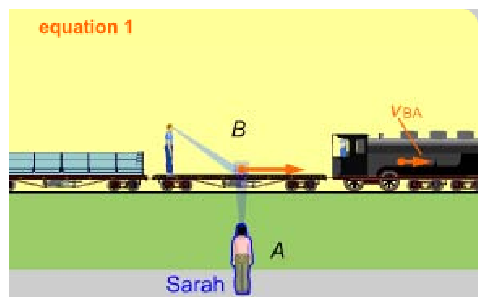
The answer is that she sees the ball as stationary. The sum of the velocities equals zero, because it is the sum of +5 m/s (the velocity of the object as measured in reference frame B) and −5 m/s (the velocity of frame B as measured from frame A).



Observer on train
Measures ball velocity relative to train



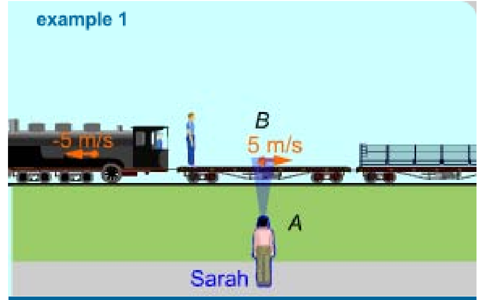
Observer on ground
Train is moving
Velocity = sum of ball, train velocities



Relative velocity equation

$$\mathbf{v}_{\mathrm{OA}} = \mathbf{v}_{\mathrm{OB}} + \mathbf{v}_{\mathrm{BA}}$$

$\mathbf{v}_{\mathrm{OA}}$ = velocity of object measured in reference frame A
$\mathbf{v}_{\mathrm{OB}}$ = velocity of object measured in reference frame B
$\mathbf{v}_{\mathrm{BA}}$ = velocity of frame B measured in frame A

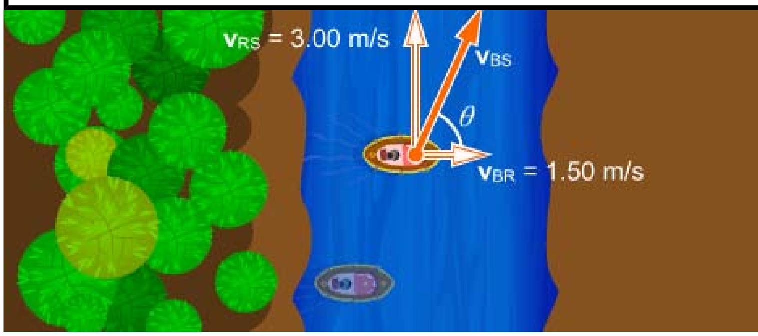


Sarah observes the train moving to the left at −5 m/s. Fred, on the train, sees the ball moving to the right at +5 m/s. What is the ball's velocity in Sarah's reference frame?

$\mathbf{v}_{OA} = \mathbf{v}_{OB} + \mathbf{v}_{BA}$

$\mathbf{v}_{OA} = (5\text{ m/s}) + (-5\text{ m/s}) = 0\text{ m/s}$

4.23 - Sample problem: relative velocity



A boat's engine propels it horizontally at 1.50 m/s, while the river current flows at 3.00 m/s. What is the velocity of the boat relative to the shore?

Variables

velocity of boat relative to shore	$\mathbf{v}_{BS}$
velocity of boat relative to river	$\mathbf{v}_{BR}$ = 1.50 m/s, in the $+x$ direction
velocity of river relative to shore	$\mathbf{v}_{RS}$ = 3.00 m/s, in the $+y$ direction

What is the strategy?

1. Identify the two reference frames, the river and the shore.
2. Write a relative velocity equation that relates the velocity of the boat relative to the shore to the two known velocities: the velocity of the boat relative to the river and the velocity of the river frame relative to the shore frame.
3. Add the velocity vectors using trigonometry to solve this equation.

Physics principles and equations

We use the relative velocity equation

$$\mathbf{v}_{OA} = \mathbf{v}_{OB} + \mathbf{v}_{BA}$$

Mathematics equations

We use the Pythagorean theorem

$$c^2 = a^2 + b^2$$

And the definition of cosine

$$\cos\theta = \text{adjacent leg / hypotenuse}$$

Step-by-step solution

First, we use the Pythagorean theorem to determine the magnitude of the velocity of the boat relative to the shore. The two vectors shown in the diagram above form the legs of a right triangle.

Step	Reason
1. $\mathbf{v}_{BS} = \mathbf{v}_{BR} + \mathbf{v}_{RS}$	equation stated above
2. $v_{BS}^2 = v_{BR}^2 + v_{RS}^2$	Pythagorean theorem
3. $v_{BS} = \sqrt{v_{BR}^2 + v_{RS}^2}$	square root
4. $v_{BS} = \sqrt{(1.50\ \text{m/s})^2 + (3.00\ \text{m/s})^2}$	enter values
5. $v_{BS} = 3.35$ m/s	evaluate

Since the question asked for the velocity, we need to determine the direction. We use trigonometry to determine the angle. We will use one leg and the value for the hypotenuse we just calculated.

Step	Reason
6. $\cos\theta = \dfrac{v_{BR}}{v_{BS}}$	definition of cosine
7. $\cos\theta = \dfrac{1.50\ \text{m/s}}{3.35\ \text{m/s}}$	enter values
8. $\theta = \arccos(0.448)$	take arccosine
9. $\theta = 63.4°$	evaluate

4.24 - Gotchas

A ball will land at the same time if you drop it straight down from the top of a building or if you throw it out horizontally. Yes, the ball will hit the ground at the same time in both cases. Velocity in the y direction is independent of velocity in the x direction.

An object has positive velocity along the x and y axes. Along the y axis, it accelerates, has a constant velocity for a while, then accelerates some more. What happens along the x axis? You have no idea. Information about motion along the y axis tells you nothing about motion along the x axis, because they can change independently.

A projectile has zero acceleration at its peak. No, a projectile has zero y velocity at its peak. Even though it briefly comes to rest in the vertical dimension, the projectile is always accelerating at −9.80 m/s^2 due to the force of gravity.

4.25 - Summary

Analyzing motion in multiple dimensions is similar to analyzing motion in one dimension. Vectors are essential here.

In multiple dimensions, the displacement vector $\Delta\mathbf{r}$ points from the initial to the final position.

Like any vector, velocity can be broken into its component vectors in the x and y dimensions using trigonometry. These components sum to equal the velocity vector. The components can also be analyzed separately, reducing a two-dimensional problem to two separate one-dimensional problems. The same principle applies to the acceleration vector.

Objects that move solely under the influence of gravity are called projectiles. To analyze projectile motion, consider the motion along each dimension separately.

A reference frame is a coordinate system used to make observations.

Observers in different reference frames will measure different velocities for an object if there is relative motion between their reference frames. To convert from one measurement to the other, use the relative velocity equation.

Equations

$$\Delta\mathbf{r} = \mathbf{r}_f - \mathbf{r}_i$$

$$\bar{\mathbf{v}} = \frac{\Delta\mathbf{r}}{\Delta t}$$

$$\bar{\mathbf{a}} = \frac{\Delta\mathbf{v}}{\Delta t}$$

Range equation

$$\Delta x = \frac{-v^2 \sin 2\theta}{a_y}$$

Shooting angle equation

$$\theta = \tfrac{1}{2}\arcsin\left(\frac{-a_y\Delta x}{v^2}\right)$$

Relative velocity equation

$$\mathbf{v}_{OA} = \mathbf{v}_{OB} + \mathbf{v}_{BA}$$

Chapter 4 Problems

Conceptual Problems

C.1 How can you change only one component of the velocity without changing the speed?

C.2 Two girls decide to jump off a diving board. Katherine steps off the diving board. Anna runs straight off the diving board so that her initial velocity is solely horizontal. They both leave the diving board at the same time. Which one lands in the water first?

i. Katherine
ii. Anna
iii. They land at the same time

C.3 An astronaut throws a baseball horizontally with the same initial velocity: once on the Earth, and once on the Moon. Where does the ball travel farther? Explain your answer.

i. On the Earth
ii. On the Moon
iii. The same distance

C.4 What is the name for the shape of a projectile's path?

C.5 Sarah is standing on the east bank of a river that runs south to north, and she wants to meet her friends who are on a boat in the middle of the river, directly west of her. Her friends point the boat east, directly at her, and then turn on the motor. The boat stays pointed east throughout its journey. (a) At what speed should Sarah walk or run along the bank of the river in order for her to meet the boat when it touches the bank? (b) In what direction should she move? (c) Does either answer depend on the speed of the boat relative to the water, which is provided by the motor?

(a) i. The speed of the river
ii. Zero
iii. Twice the speed of the river
iv. The speed that the boat could travel in still water

(b) i. North
ii. South
iii. She should stand still

(c) ○ Yes ○ No

C.6 Two students, Jim and Sarah, are walking to different classrooms from the same cafeteria. How does Jim's velocity in Sarah's reference frame relate to Sarah's velocity in Jim's reference frame?

i. There is no relationship
ii. They are identical
iii. They are equal but opposite
iv. There is not enough information to answer

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the driving interactive problem from this section to answer the following questions. (a) Does changing the *x* velocity affect the *y* velocity? (b) How does increasing the magnitude of either velocity affect the overall speed?

(a) ○ Yes ○ No

(b) i. There is no effect
ii. The speed increases
iii. The speed decreases

0.2 Use the simulation in the cannon interactive in this section to answer the following questions. (a) Does the cannonball's horizontal velocity change as it moves through the air? (b) Does its vertical velocity change? (c) If you increase the cannonball's horizontal velocity, does it stay in the air longer, the same amount of time, or shorter? (d) Who is inside the left haystack?

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c) i. Longer
ii. The same
iii. Shorter
(d) i. Elmo
ii. Elvis
iii. Marie Curie

Section 1 - Displacement in two dimensions

1.1 Calculate the displacement of a ladybug that starts at (12.0, 5.0, 5.0) m and ends at (3.5, 2.0, 4.0) m.

(__________ , __________ , __________) m

1.2 In a visit to the nation's capital, a foreign head of state travels from (−3.0, −4.0) km to (4.0, 7.0) km. She then travels with a displacement of (−5.0, −3.9) km. Calculate the her total displacement over the course of the trip.

(__________ , __________) km

1.3 What is the displacement from a starting position of (14.0, 3.0) m to a final position of (−3.0, −4.0) m?

(__________ , __________) m

1.4 An irresolute boater drives south at 28 km/h for a half hour, pauses for 5 minutes, and then drives east for 45 minutes at 32 km/h. State his displacement vector in kilometers using rectangular coordinates. Use the convention that north and east are positive.

(__________ , __________) km

1.5 Sara's golf ball is 5 meters south and 4 meters west of a hole. What is the displacement vector from the ball's current position to the hole in rectangular coordinates? Assume that north and east are in the positive direction.

(__________ , __________) m

Section 2 - Velocity in two dimensions

2.1 A golf ball is launched at a 37.0° angle from the horizontal at an initial velocity of 48.6 m/s. State its initial velocity in rectangular coordinates.

(__________ , __________) m/s

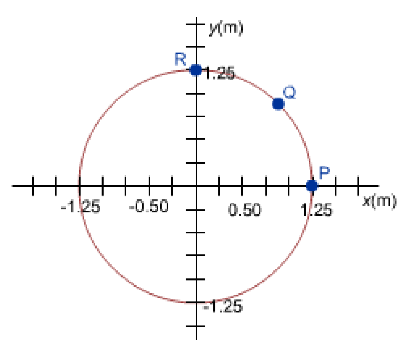
2.2 The displacement vector for a 15.0 second interval of a jet airplane's flight is (2.15e+3, −2430) m. (a) What is the magnitude of the average velocity? (b) At what angle, measured from the positive *x* axis, did the airplane fly during this time interval?

(a) __________ m/s
(b) __________ °

2.3 An airplane is flying horizontally at 115 m/s, at a 20.5 degree angle north of east. State its velocity in rectangular coordinates. Assume that north and east are in the positive direction.

(__________ , __________) m/s

2.4 A particle travels counterclockwise around the origin at constant speed in a circular path that has a diameter of 2.50 m, going around twice per second. Point Q is on the path halfway between points P and R, which lie on the *x* axis and the *y* axis, respectively. State the following velocites as vectors in polar notation. (a) What is the particle's average velocity over the interval PR? (b) What is its average velocity over the interval PQ? (c) What is its instantaneous velocity at point P?



(a) __________ m/s, __________ °
(b) __________ m/s, __________ °
(c) __________ m/s, __________ °

2.5 The Earth moves in a roughly circular orbit around the Sun, at an average distance from the Sun of 1.49589×10^{11} meters over a period of 365.24 days. The coordinate system is fixed, with the sun at the origin. (a) What is the magnitude of the Earth's displacement in one day? (b) If you calculate the magnitude of the displacement of the Earth in one day by finding its average velocity over the course of an hour and then multiplying by 24 hours, would this new displacement be the same as, larger than, or smaller than that calculated in part a? (c) If you calculated the Earth's average velocity over a day instead of an hour would the magnitude of its average velocity be greater or lesser? (d) Explain your answer to part (c).

(a) ________ m

(b) i. Same
ii. Larger
iii. Smaller

(c) i. Same
ii. Larger
iii. Smaller

(d) Submit answer on paper

Section 4 - Acceleration in two dimensions

4.1 An ice skater starts out traveling with a velocity of (−3.0, −7.0) m/s. He performs a 3.0 second maneuver and ends with a velocity of (0, 5.0) m/s. (a) What is his average acceleration over this period? (b) A different ice skater starts with the same initial velocity, accelerates at (1.5, 3.5) m/s^2 for 2.0 seconds, and then at (0, 5.0) m/s^2 for 1.0 seconds. What is his final velocity?

(a) (________ , ________) m/s 2
(b) (________ , ________) m/s

4.2 A spaceship accelerates at (6.30, 3.50) m/s^2 for 6.40 s. Given that its final velocity is (450, 570) m/s, find its initial velocity.

(________ , ________) m/s

4.3 An airplane is flying with velocity (152, 0) m/s. It is accelerated in the *x* direction by its propeller at 8.92 m/s^2 and in the negative *y* direction by a strong downdraft at 1.21 m/s^2. What is the airplane's velocity after 3.10 seconds?

(________ , ________) m/s

4.4 An electron starts at rest. Gravity accelerates the electron in the negative *y* direction at 9.80 m/s^2 while an electric field accelerates it in the positive *x* direction at 3.80×10^6 m/s^2. Find its velocity 2.45 s after it starts to move.

(________ , ________) m/s

4.5 A particle's initial velocity was (3.2, 4.5) m/s, its final velocity is (2.6, 5.8) m/s, and its constant acceleration was (−2.4, 5.2) m/s^2. What was its displacement?

(________ , ________) m

Section 7 - Projectile motion

7.1 A friend throws a baseball horizontally. He releases it at a height of 2.0 m and it lands 21 m from his front foot, which is directly below the point at which he released the baseball. (a) How long was it in the air? (b) How fast did he throw it?

(a) ________ s
(b) ________ m/s

7.2 A cannon mounted on a pirate ship fires a cannonball at 125 m/s horizontally, at a height of 17.5 m above the ocean surface. Ignore air resistance. (a) How much time elapses until it splashes into the water? (b) How far from the ship does it land?

(a) ________ s
(b) ________ m

7.3 A juggler throws a ball straight up into the air and it takes 1.20 seconds to reach its peak. How many seconds will it take after that to fall back into his hand? Assume he throws and catches the ball at the same height.

________ s

7.4 A juggler throws a ball with an initial horizontal velocity of +1.1 m/s and an initial vertical velocity of +5.7 m/s. What is its acceleration at the top of its flight path? Make sure to consider the sign when responding. Consider the upward direction as positive.

________ m/s 2

7.5 The muzzle velocity of an armor-piercing round fired from an M1A1 tank is 1770 m/s (nearly 4000 mph or mach 5.2). A tank is at the top of a cliff and fires a shell horizontally. If the shell lands 6520 m from the base of the cliff, how high is the cliff?

________ m

7.6 A cannon is fired horizontally from atop a 40.0 m tower. The cannonball travels 145 m horizontally before it strikes the ground. With what velocity did the ball leave the muzzle?

__________ m/s

7.7 A box is dropped from a spacecraft moving horizontally at 27.0 m/s at a distance of 155 m above the surface of a moon. The rate of freefall acceleration on this airless moon is 2.79 m/s^2. (a) How long does it take for the box to reach the moon's surface? (b) What is its horizontal displacement during this time? (c) What is its vertical velocity when it strikes the surface? (d) At what speed does the box strike the moon?

(a) __________ s
(b) __________ m
(c) __________ m/s
(d) __________ m/s

7.8 Two identical cannons fire cannonballs horizontally with the same initial velocity. One cannon is located on a platform partway up a tower, and the other is on top, four times higher than the first cannon. How much farther from the tower does the cannonball from the higher cannon land than the cannonball from the lower cannon?

i. 1 times farther
ii. 2
iii. 3
iv. 4
v. 16

7.9 A ball rolls at a constant speed on a level table a height h above the ground. It flies off the table with a horizontal velocity v and strikes the ground. Choose the correct expression for the horizontal distance it travels, as measured from the table's edge.

(a) $v\left(-\sqrt{\frac{2h}{g}}\right)$

(b) $v\sqrt{\frac{2h}{g}}$

(c) $v\sqrt{-\frac{2g}{h}}$

(d) $v\sqrt{\frac{h}{-2g}}$

i. a
ii. b
iii. c
iv. d

7.10 An archer wants to determine the speed at which her new bow can fire an arrow. She paces off a distance d from her target and fires her arrow horizontally. She determines that she launched it at a height h higher than where it strikes the target. She ignores air resistance. Choose an expression in terms of h, d and g that she can use to determine the initial speed of her arrow.

(a) $d\sqrt{\frac{g}{2h}}$

(b) $d\sqrt{\frac{h}{g}}$

(c) $d\sqrt{\frac{-g}{h}}$

i. a
ii. b
iii. c

7.11 Randy Johnson throws a baseball horizontally from the top of a building as fast as his 2004 record of 102.0 mph (45.6 m/s). How much time passes until it moves at an angle 13.0 degrees below the horizontal? Ignore air resistance.

__________ s

7.12 Sam tosses a ball horizontally off a footbridge at 3.1 m/s. How much time passes after he releases it until its speed doubles?

__________ s

7.13 You are in a water fight, and your tricky opponent, wishing to stay dry, goes into a nearby building to spray you from above. His first attempt falls one-third of the way from the base of the building to your position. He climbs up to a higher window and shoots again, and manages to give you a good soaking. In both cases, he aims his water gun parallel to the ground. What is the ratio of his first height to his second height?

i. 1 to 1
ii. 1 to 2
iii. 1 to 3
iv. 1 to 4
v. 1 to 9
vi. 1 to 11

7.14 You fire a squirt gun horizontally from an open window in a multistory building and make note of where the spray hits the ground. Then you walk up to a window 5.0 m higher and fire the squirt gun again, discovering that the water goes 1.5 times as far. Ignore air resistance. How long does the second shot take to hit the ground?

__________ s

7.15 Two cannonballs are fired horizontally at a velocity *v*. One is fired from a height of *h*, and one 3*h*. (a) State an expression for the ratio of the displacement of the cannonball fired at 3*h* to the displacement of the one fired at *h* before either hits the ground. (b) State an expression, in terms of *v* and *h*, for the ratio of the displacement from the firing point to the landing point of the cannonball fired at 3*h* to the one fired at *h*. To be clear: the displacement is the "total" displacement, both vertical and horizontal.

(a) i. 1 to 1
ii. 1 to 2
iii. 1 to 3
iv. 2 to 1
v. 3 to 1

(b) Submit answer on paper.

Section 9 - Interactive problem: the monkey and the professor

9.1 Using the simulation in the interactive problem in this section, where should the monkey aim the bazooka? Why?

i. Above the professor's glove
ii. At the professor's glove
iii. Below the professor's glove

Section 11 - Projectile motion: juggling

11.1 A juggler throws a ball from height of 0.950 m with a vertical velocity of + 4.25 m/s and misses it on the way down. What is its velocity when it hits the ground?

__________ m/s

11.2 You are returning a tennis ball that your five-year-old neighbor has accidentally thrown into your yard, but you need to throw it over a 3.0 m fence. You want to throw it so that it barely clears the fence. You are standing 0.50 m away from the fence and throw underhanded, releasing the ball at a height of 1.0 m. State the initial velocity vector you would use to accomplish this.

(__________ , __________) m/s

11.3 Your friend has climbed a tree to a height of 6.00 m. You throw a ball vertically up to her and it is traveling at 5.00 m/s when it reaches her. What was the speed of the ball when it left your hand if you released it at a height of 1.10 m?

__________ m/s

11.4 Two people perform the same experiment twice, once on Earth and once on another planet. One observes from the top of a 40.5 m cliff, and the other stands at ground level. The person at the bottom throws a ball at a speed of 31.0 m/s at a 75.4° angle above the horizontal. (a) When this experiment is being conducted on Earth, what **vertical** velocity would you expect the ball to have when it reaches a height of 40.5 m? (b) On the other planet, the ball is observed to have a vertical velocity of 24.5 m/s when it reaches that height. Assuming there are no forces besides gravity acting on the ball, what is the acceleration due to gravity on the other planet? State your answer as a positive value, just as *g* is stated.

(a) __________ m/s
(b) __________ m/s 2

11.5 A friend of yours has climbed a tree to a height of 12.0 m, but he forgot his bag lunch and is very hungry. You cannot toss it straight up to him because there are branches in the way. Instead, you throw the bag from a height of 1.20 m, at a horizontal distance of 3.50 m from the location of your friend. If you want the bag to have zero vertical velocity when it reaches your friend, what must the horizontal and vertical velocity components of your throw be? Assume you are throwing the bag in the positive horizontal direction.

(__________ , __________) m/s

Section 13 - Interactive problem: the monkey and the professor, part II

13.1 Using the simulation in the interactive problem in this section, where should the monkey aim the bazooka? Why?

i. Above the professor's glove
ii. At the professor's glove
iii. Below the professor's glove

Section 14 - Projectile motion: aiming a cannon

14.1 A long jumper travels 8.95 meters during a jump. He moves at 10.8 m/s when starts his leap. At what angle from the horizontal must he have been moving when he started his jump? You may need the double-angle formula:
2 sin *u* cos *u* = sin (2*u*)

__________ °

14.2 Your projectile launching system is partially jammed. It can only launch objects with an initial vertical velocity of 42.0 m/s, though the horizontal component of the velocity can vary. You need your projectile to land 211 m from its launch point. What horizontal velocity do you need to program into the system?

__________ m/s

14.3 A professional punter can punt a ball so that its "hang time" is 4.20 seconds and it travels 39.0 m horizontally. State its initial velocity in rectangular coordinates. Ignore air resistance, and assume the punt receiver catches the ball at the same height at which the punter kicks it.

(__________ , __________) m/s

14.4 Consider the second monkey and the professor simulation in the textbook. Use θ for the angle at which the cannon is aimed, and *v* for the initial speed of the banana. Show mathematically why θ must be the angle such that the bazooka initially points at the professor in order for the professor to catch the banana.

14.5 A home run just clears a fence 105 m from home plate. The fence is 4.00 m higher than the height at which the batter struck the ball, and the ball left the bat at a 31.0° angle above the horizontal. At what speed did the ball leave the bat?

__________ m/s

14.6 You are the stunt director for a testosterone-laden action movie. A car drives up a ramp inclined at 10.0° above the horizontal, reaching a speed of 40.0 m/s at the end. It will jump a canyon that is 101 meters wide. The lip of the takeoff ramp is 201 meters above the floor of the canyon. (a) How long will the car take to cross the canyon? (b) What is the maximum height of the cliff on the other side so that the car lands safely? (c) What angle with the horizontal will the car's velocity make when it lands on the other side? Assume the height of the other side is the maximum value you just calculated. Express the angle as a number between −180° and +180°.

(a) __________ s
(b) __________ m
(c) __________ °

14.7 The infamous German "Paris gun" was used to launch a projectile with a flight time of 170 s for a horizontal distance of 122 km. Based on this information, and ignoring air resistance as well as the curvature of the Earth, calculate (a) the gun's muzzle speed and (b) the angle, measured above the horizontal, at which it was fired.

(a) __________ m/s
(b) __________ °

14.8 At its farthest point, the three-point line is 7.24 meters away from the basket in the NBA. A basketball player stands at this point and releases his shot from a height of 2.05 meters at a 35.0 degree angle. The basket is 3.05 meters off the ground. The player wants the ball to go directly in (no bank shots). At what speed should he throw the ball?

__________ m/s

14.9 You and your friend are practicing pass plays with a football. You throw the football at a 35.0° angle above the horizontal at 19.0 m/s. Your friend starts right next to you, and he moves down the field directly away from you at 5.50 m/s. How long after he starts running should you throw the football? Assume he catches the ball at the same height at which you throw it.

__________ s

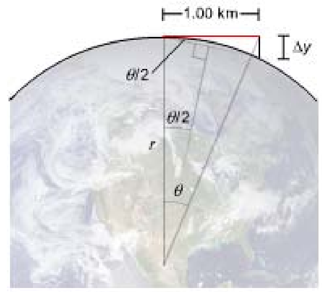
14.10 If there were no air resistance, how fast would you have to throw a football at an initial 45° angle in order to complete an 80 meter pass?

__________ m/s

14.11 A soccer ball is lofted toward the goal from a distance of 9.00 m. It has an initial velocity of 12.7 m/s and is kicked at an angle of 72.0 degrees above the horizontal. (a) When the ball crosses the goal line, how high will it be above the ground? (b) How long does it take the ball to reach the goal line? (It will still be in the air.)

(a) __________ m
(b) __________ s

14.12 In this problem, you will be asked to analyze portions of a famous "thought experiment." To explain the orbital motion of the Moon, Newton imagined a cannon that could fire a cannonball horizontally at a great speed. He reasoned that the cannonball's motion was fundamentally the same as the Moon's. He linked projectile motion with orbits, bringing celestial bodies into the realm of everyday experience.



Consider the Earth to be a perfect sphere of radius $r = 6.38\times10^6$ m. Imagine that the cannon is mounted on a tower 125 meters above the surface of the Earth. Ignore, as Newton did, air resistance, mountains and the motion of the Earth.

The cannon fires the cannonball horizontally, tangent to the Earth's surface. The faster the cannon fires the ball, the farther it goes before striking the ground. You wish to fire a ball at a speed so great that it never hits the ground: The spherical Earth curves down and away from it at exactly the same rate at which it falls. In this scenario the cannonball will describe a circular trajectory, circling the Earth at a constant height of 125 m. It will return to strike the cannon in the rear.

To determine the necessary speed, consider what happens during the first 1.00 km of horizontal displacement of the cannonball. You will calculate the vertical displacement, using the curvature of the Earth.

(a) Suppose you draw a line tangent to the Earth's surface at the location of the cannon tower, and extend it exactly 1.00 km in the direction of the cannonball's path. This is shown, with some exaggeration, with the red line in the diagram. What is the angle θ (in radians) that this line segment subtends at the center of the Earth? (b) You can draw a right-angled line from the end of the red segment down to the ground in order to measure the "fall-away" Δy of the Earth's surface. This very nearly equals the opposite side of a triangle whose base angle is $\theta/2$. Use tan $\theta/2$ in an equation to approximate the Earth's vertical fall-away Δy in a horizontal distance of 1.00 km (state your answer as a positive number). (c) How fast must you fire the cannonball so that its fall after one horizontal kilometer equals the fall-away of the Earth? (d) Newton used the cannon as a "thought experiment" to explain (circular) orbital motion. Why is the trajectory of the cannonball **not** parabolic?

(a) __________ rad
(b) __________ m
(c) __________ m/s
(d) Submit answer on paper

14.13 Achilles is lurking inside a high wall that defends an ancient city. At his location, it is 10.0 m thick and 25.0 m high. An enemy archer wishes to hit Achilles with his arrow. He can fire his arrows at 50.0 m/s, and he is crouching 27.0 m from the outside of the city wall. The release point of the arrow and its landing height are both at ground level. (a) If Achilles stands up against the inside surface of the wall he will be safe. If he begins to walk away from the wall, at what distance from it will the archer be able to hit him? (b) Xena is inside the city, very far from the wall, and she is also safe. Suppose Xena starts walking toward the wall. At what distance from the inside surface of the wall will she no longer be safe?

(a) __________ m
(b) __________ m

14.14 The circus has come to town, and Lael the Magnificent will kick a grape into her assistant's waiting mouth (not to worry − the grape is freshly washed). She starts at the origin and boots the grape at an angle 20.0° east of north. The grape starts from a height of 0.200 meters, with a velocity of 12.0 m/s, 30.0° above the horizontal. Her assistant races to intercept the grape as it descends. Assume that his mouth is 1.75 meters above the ground. (a) How much time elapses after the kick until he catches the grape? (b) At what position must he be standing for the grape to fall into his mouth? Use the *x* dimension for east-west, and the *y* dimension for north-south.

(a) __________ s
(b) (__________ , __________) m

Section 18 - Interactive problem: the human cannonball

18.1 Using the simulation in the interactive problem in this section, if the cannon is aimed at an angle of 20.0°, what should the firing speed be so that the human cannonball will land on the net?

__________ m/s

Section 19 - Interactive problem: test your juggling!

19.1 If the juggler in the simulation in the interactive problem in this section wants to throw the balls with a horizontal velocity of 1.1 m/s, what should the vertical velocity be?

__________ m/s

19.2 If the juggler in the simulation in the interactive problem in this section wants the balls to remain in the air for 1.4 seconds, (a) what should the vertical velocity be? (b) What should the horizontal velocity be?

(a) __________ m/s
(b) __________ m/s

Section 20 - The range and elevation equations

20.1 You have a cannon with a muzzle velocity of 155 m/s and a target 1950 m away. State (a) the smaller angle and (b) the larger angle at which you can aim the cannon and reach the target.

(a) The smaller angle __________ °
(b) The greater angle __________ °

20.2 You desperately want to qualify for the Olympics in the long jump, so you decide to hold the qualifying event on the moon of your choice. You need to jump 7.52 m (and conveniently beat Galina Chistyakova's record) to qualify. The maximum speed at which you can run at any location is 5.90 m/s. What is the magnitude of the maximum rate of freefall acceleration the moon can have for you to achieve your dream?

__________ m/s 2

20.3 Natalya Lisovskaya set a track and field record in 1987 when she put the shot so that its horizontal displacement was 22.63 m. Assume that the angle of the initial velocity of the shot was such that it maximized the range, and for simplicity, assume the height of the ball when it landed was the same as the height at which it left her hand. If she were to put the shot vertically with the same speed, what is the maximum height above her hand that the shot would reach?

__________ m

20.4 You want to throw a ball to your friend so that it has the least speed possible when she catches it. If the throw and the catch occur at the same height, at what angle should you throw the ball?

__________ °

20.5 You kick a soccer ball at a 10.0 degree angle above the horizontal. You want to double how far it travels but still want to kick the ball with equal initial speed. You cannot kick the ball above a 45 degree angle. At what angle should you kick the ball?

__________ °

Section 22 - Relative velocity

22.1 Agent Bond is in the middle of one of his trademark, nearly impossible getaways. He is in a convertible driving west at a speed of 23.0 m/s (this is the reading on his speedometer) on top of a train, heading toward the back end. The train is moving horizontally, due east at 15.0 m/s. As the convertible goes over the edge of the last train car, Bond jumps off. With what horizontal velocity relative to the convertible should he jump to hit the ground with a horizontal velocity of 0 m/s, so that he merely shakes, but does not spill (nor stir) his drink? Consider east to be the positive direction.

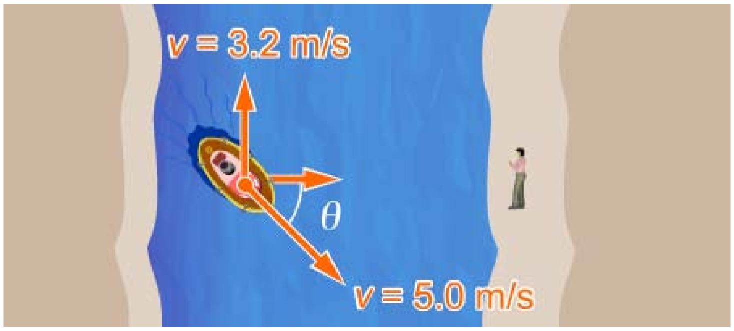
__________ m/s

22.2 A boat is moving on Lake Tahoe with a velocity (6.5, 9.5) m/s relative to the water. However, a strong wind is causing the water on the surface to move with velocity (−4.5, −2.5) m/s relative to the land, and this velocity must be added to the boat's. What is the velocity of the boat, as seen from a skateboarder on the shore with a velocity of (3.5, −0.70) m/s relative to the land?

(__________ , __________) m/s

Section 23 - Sample problem: relative velocity

23.1 Sarah's friends have decided that it is inconsiderate to make Sarah walk down the riverbank to meet them, so they decide to come straight across the river to meet her. The motor propels the boat at a speed of 5.0 m/s relative to the water and the current flows at 3.2 m/s with respect to the ground. If they want to move directly across the river without being deflected downstream or moving upstream, find the angle between the direction in which they need to point the boat and their path across the river. State the angle as positive.

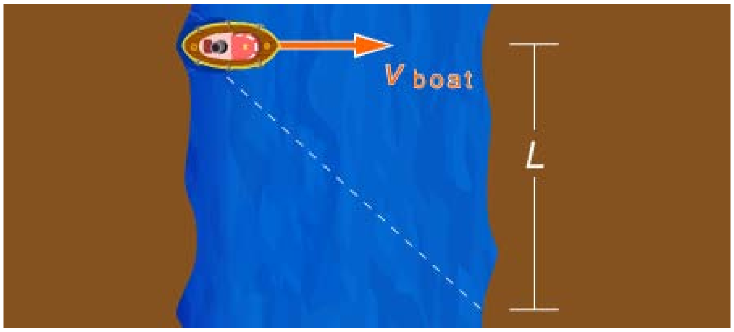


__________ °

23.2 A plane is pointed in the positive x direction and is flying with an speed of (200, 0, 0) m/s relative to the air. However, due to an exceptionally strong updraft, the air is moving in the positive z direction at 50.0 m/s with respect to the ground. Find the velocity of the plane with respect to the ground.

(__________ , __________ , __________) m/s

23.3 You have a boat with a motor that propels it at v_{boat} = 4.5 m/s relative to the water. You point it directly across the river and find that when you reach the other side, you have traveled a total distance of 27 m (indicated by the dotted line in the diagram) and wound up 12 m downstream. What is the speed of the current?



__________ m/s

23.4 You are riding a bike at 15 m/s due east, the positive x direction. You throw a tennis ball horizontally toward the north (positive y direction) at 10 m/s in your reference frame. On the tennis ball is a fly that takes off vertically from the ball (positive z direction) at 3.0 m/s in the ball's reference frame as soon as it leaves your hand. The air moves with velocity (−2.0, 2.0, 0.0) m/s with respect to the ground due to a gentle breeze. Find the velocity of the fly with respect to the air.

(__________ , __________ , __________) m/s

23.5 A softball player is running towards first base with a velocity of (10.0, 0.0) m/s. The shortstop is running towards third base with a velocity of (−5.0, 0.0) m/s. A fly ball is moving with a velocity of (6.0, −18.0) m/s. Find the velocities of the runner, the shortstop, and the ball when viewed from the reference frame of each of the others.

(a) Velocity of shortstop in runner's reference frame
(__________ , __________) m/s
(b) Velocity of baseball in runner's reference frame
(__________ , __________) m/s
(c) Velocity of runner in shortstop's reference frame
(__________ , __________) m/s
(d) Velocity of baseball in shortstop's reference frame
(__________ , __________) m/s
(e) Velocity of runner in baseball's reference frame
(__________ , __________) m/s
(f) Velocity of shortstop in baseball's reference frame
(__________ , __________) m/s

Chapter 5 Force and Newton's Laws

5.0 - Introduction

Objects can speed up, slow down, and change direction while they move. In short, they accelerate.

A famous scientist, Sir Isaac Newton, wondered how and why this occurs. Theories about acceleration existed, but Newton did not find them very convincing. His skepticism led him to some of the most important discoveries in physics.

Before Newton, people who studied motion noted that the objects they observed on Earth always slowed down. According to their theories, objects possessed an internal property that caused this acceleration. This belief led them to theorize that a force was required to keep things moving.

This idea seems like common sense. Moving objects do seem to slow down on their own: a car coasts to a stop, a yo-yo stops spinning, a soccer ball rolls to a halt. Newton, however, rejected this belief, instead suggesting the opposite: The nature of objects is to continue moving unless some force acts on them. For instance, Newton would say that a soccer ball stops rolling because of forces like friction and air resistance, not because of some property of the soccer ball. He would say that if these forces were **not** present, the ball would roll and roll and roll. A force (a kick) is required to start the ball's motion, and a force such as the frictional force of the grass is required to stop its motion.

Newton proposed several fundamental principles that govern forces and motion. Nearly 300 years later, his insights remain the foundation for the study of forces and much of motion. This chapter stands as a testament to a brilliant scientist.

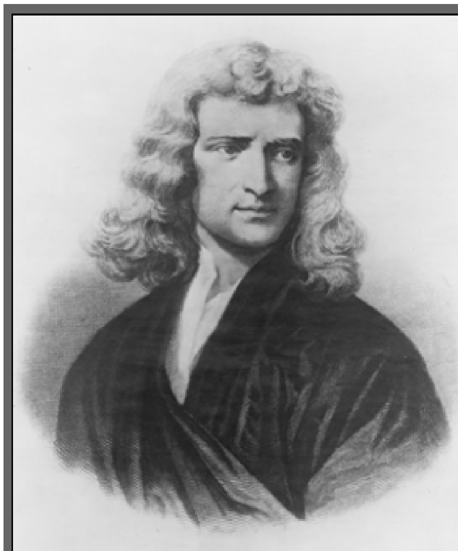
At the right, you can use a simulation to experience one of Newton's fundamental principles: his law relating a net force, mass and acceleration. In the simulation, you can attempt some of the basic tasks required of a helicopter pilot. To do so, you control the **net** force upward on the helicopter. When the helicopter is in the air, the net force equals the lift force minus its weight. (The lift force is caused by the interaction of the spinning blades with the air, and is used to propel the helicopter upward.) The net force, like all forces, is measured in newtons (N).

When the helicopter is in the air, you can set the net force to positive, negative, or zero values. The net force is negative when the helicopter's lift force is less than its weight. When the helicopter is on the ground, there cannot be a negative net force because the ground opposes the downward force of the helicopter's weight and does not allow the helicopter to sink below the Earth's surface.

The simulation starts with the helicopter on the ground and a net force of 0 N. To increase the net force on the helicopter, press the up arrow key (↑) on your keyboard; to decrease it, press the down arrow key (↓). This net force will continue to be applied until you change it.

interactive 1

Force, mass and acceleration
Lift the helicopter off the ground



Sir Isaac Newton, 1642 -1727

To start, apply a positive net force to cause the helicopter to rise off the ground. Next, attempt to have the helicopter reach a constant vertical velocity. For an optional challenge, have it hover at a constant height of 15 meters, and finally, attempt to land (not crash) the helicopter.

Once in the air, you may find that controlling the craft is a little trickier than you anticipated − it may act a little skittish. Welcome to (a) the challenge of flying a helicopter and (b) Newton's world.

Here are a few hints: Start slowly! Initially, just use small net forces. You can look at the acceleration gauge to see in which direction you are accelerating. Try to keep your acceleration initially between plus or minus 0.25 m/s^2.

This simulation is designed to help you experiment with the relationship between force and acceleration. If you find that achieving a constant velocity or otherwise controlling the helicopter is challenging − read on! You will gain insights as you do.

5.1 - Force

Force: Loosely defined as "pushing" or "pulling."

Your everyday conception of force as pushing or pulling provides a good starting point for explaining what a force is.

There are many types of forces. Your initial thoughts may be of forces that require direct contact: pushing a box, hitting a ball, pulling a wagon, and so on.

Some forces, however, can act without direct contact. For example, the gravitational force of the Earth pulls on the Moon even though hundreds of thousands of kilometers separate the two bodies. The gravitational force of the Moon, in turn, pulls on the Earth.

Electromagnetic forces also do not require direct contact. For instance, two magnets will attract or repel each other even when they are not touching each other.

We have discussed a few forces above, and could continue to discuss more of them: static friction, kinetic friction, weight, air resistance, electrostatic force, tension, buoyant force, and so forth. This extensive list gives you a sense of why a general definition of force is helpful.

These varied types of forces do share some essential attributes. Newton observed that a force, or to be precise, a net force, causes acceleration.

All forces are vectors: their direction matters. The weightlifter shown in Concept 1 must exert an upward force on the barbell in order to accelerate it off the ground. For the barbell to accelerate upward, the force he exerts must be greater than the downward force of the Earth's gravity on the barbell. The *net force* (the vector sum of all forces on an object) and the object's mass determine the direction and amount of acceleration.

The SI unit for force is the newton (N). One newton is defined as one kg·m/s^2. We will discuss why this combination of units equals a newton shortly.

We have given examples where a net force causes an object to accelerate. Forces can also be in equilibrium (balance), which means there is no net force and no acceleration. When a weightlifter holds a barbell steady over his head after lifting it, his upward force on the barbell exactly balances the downward gravitational force on it, and the barbell's acceleration is zero. The net force would also be zero if he were lifting the barbell at constant velocity.

concept 1

Force
"Pushing" or "pulling"
Net force = vector sum of forces
Measured in newtons (N)
1 N = 1 kg·m/s^2

5.2 - Newton's first law

Newton's first law: "Every body perseveres in its state of being at rest or of moving uniformly straight forward except insofar as it is compelled to change by forces impressed."

This translation of Newton's original definition (Newton wrote it in Latin) may seem antiquated, but it does state an admirable amount of physics in a single sentence. Today, we are more likely to summarize Newton's first law as saying that **an object remains at rest, or maintains a constant velocity, unless a net external force acts upon it**. (Newton's formulation even includes an "insofar" to foreshadow his second law, which we will discuss shortly.)

To state his law another way: An object's velocity changes − it accelerates − when a net force acts upon it. In Concept 1, a puck is shown gliding across the ice with nearly constant velocity because there is little net force acting upon it. The puck that is stationary in Concept 2 will not move until it is struck by the hockey stick.

The hockey stick can cause a great change in the puck's velocity: a professional's slap shot can travel 150 km/hr. Forces also cause things to slow down. As a society, we spend a fair amount of effort trying to minimize these forces. For example, the grass of a soccer field is specially cut to reduce the force of friction to ensure that the ball travels a good distance when passed or shot.

Top athletes also know how to reduce air resistance. Tour de France cyclists often bike single file. The riders who follow the leader encounter less air resistance. Similarly, downhill ski racers "tuck" their bodies into low, rounded shapes to reduce air resistance, and they coat the bottoms of their skis with wax compounds to reduce the slowing effect of the snow's friction.

Newton's first law states that an object will continue to move "uniformly straight" unless acted upon by a force. Today we state this as "constant velocity," since a change in direction is acceleration as much as a change in speed. In either formulation, the point is this: Direction matters. An object not only continues at the same speed, it also moves in the same direction unless a net force acts upon it.

concept 1

Newton's first law
Objects move at constant velocity unless acted on by net force

concept 2

Newton's first law
Objects at rest remain at rest in absence of net force

You use this principle every day. Even in as basic a task as writing a note, your fingers apply changing forces to alter the direction of the pen's motion even as its speed is approximately constant.

There is an important fact to note here: Newton's laws hold true in an *inertial reference frame*. An object that experiences no net force in an inertial reference frame moves at a constant velocity. Since we assume that observations are made in such a reference frame, we will be terse here about what is meant. The surface of the Earth (including your physics lab) approximates an inertial reference frame, certainly closely enough for the typical classroom lab experiment. (The motion of the Earth makes it less than perfect.)

A car rounding a curve provides an example of a *non-inertial reference frame*. If you decided to conduct your experiments inside such a car, Newton's laws would **not** apply. Objects might seem to accelerate (a coffee cup sliding along the dashboard, for example) yet you would observe no net force acting on the cup. However, the nature of observations made in an accelerating reference frame is a topic far removed from this chapter's focus, and this marks the end of our discussion of reference frames in this chapter.

5.3 - Mass

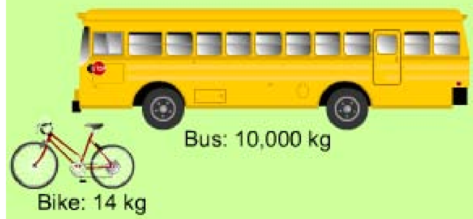
Mass: A property of an object that determines how much it will resist a change in velocity.

Newton's second law summarizes the relationship of force, mass and acceleration. Mass is crucial to understanding the second law because an object's mass determines how much it resists a change in velocity.

More massive objects require more net force to accelerate than less massive objects. An object's resistance to a change in velocity is called its *inertial mass.* It requires more force to accelerate the bus on the right at, say, five m/s^2 than the much less massive bicycle.

A common error is to confuse mass and weight. Weight is a force caused by gravity and is measured in newtons. Mass is an object's resistance to change in velocity and is measured in kilograms. An object's weight can vary: Its weight is greater on Jupiter's surface than on Earth's, since Jupiter's surface gravity is stronger than Earth's. In contrast, the object's mass does not change as it moves from planet to planet. The kilogram (kg) is the SI unit of mass.

concept 1



Mass
Measures an object's resistance to change in velocity
Measured in kilograms (kg)

5.4 - Gravitational force: weight

Weight: The force of gravity on an object.

We all experience weight, the force of gravity. On Earth, by far the largest component of the gravitational force we experience comes from our own planet. To give you a sense of proportion, the Earth exerts 1600 times more gravitational force on you than does the Sun. As a practical matter, an object's weight on Earth is defined as the gravitational force the Earth exerts on it.

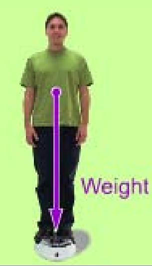
Weight is a force; it has both magnitude and direction. At the Earth' surface, the direction of the force is toward the center of the Earth.

The magnitude of weight equals the product of an object's mass and the rate of freefall acceleration due to gravity. On Earth, the rate of acceleration g due to gravity is 9.80 m/s^2. The rate of freefall acceleration depends on a planet's mass and radius, so it varies from planet to planet. On Jupiter, for instance, gravity exerts more force than on Earth, which makes for a greater value for freefall acceleration. This means you would weigh more on Jupiter's surface than on Earth's.

Scales, such as the one shown in Concept 1, are used to measure the magnitude of weight. The force of Earth's gravity pulls Kevin down and compresses a spring. This scale is calibrated to display the amount of weight in both newtons and pounds, as shown in Equation 1. Forces like weight are measured in pounds in the British system. One newton equals about 0.225 pounds.

A quick word of caution: In everyday conversation, people speak of someone who "weighs 100 kilograms," but kilograms are units for mass, not weight. Weight, like any force, is measured in newtons. A person with a mass of 100 kg weighs 980 newtons.

concept 1



Weight
Force of gravity on an object
Direction "down" (toward center of planet)

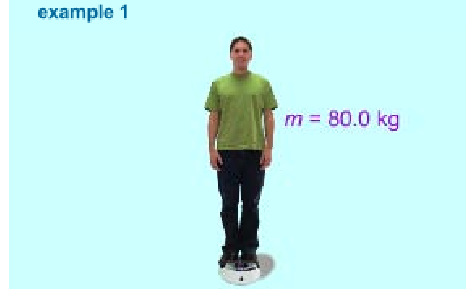
equation 1

Weight

$$W = mg$$

W = weight
m = mass
g = freefall acceleration
Units: newtons (N)

example 1



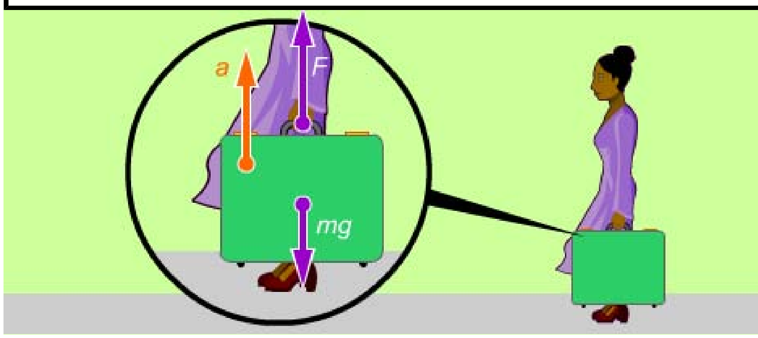
What is this person's weight on Earth?

$W = mg$

$W = (80.0\ \text{kg})(9.80\ \text{m/s}^2)$

$W = 784\ \text{N}$

5.5 - Newton’s second law



concept 1

Newton’s second law

Net force equals mass times acceleration

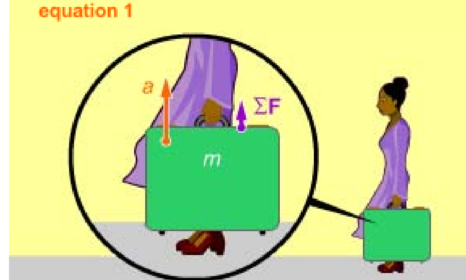
Newton's second law: “A change in motion is proportional to the motive force impressed and takes place along the straight line in which that force is impressed.”

Newton stated that a change in motion (acceleration) is proportional to force. Today, physicists call this Newton’s second law, and it is stated to explicitly include mass. Physicists state that **acceleration is proportional to the net force on an object and inversely proportional to its mass**.

To describe this in the form of an equation: net force equals mass times acceleration, or $\Sigma\mathbf{F} = m\mathbf{a}$. It is the law. The Σ notation means the vector sum of all the forces acting on an object: in other words, the net force. Both the net force and acceleration are vectors that point in the same direction, and Newton’s formulation stressed this point: “The change in motion…takes place along the straight line in which that force is impressed.” The second law explains the units that make up a newton (kg·m/s^2); they are the result of multiplying mass by acceleration.

In the illustrations, you see an example of forces and the acceleration caused by the net force. The woman who stars in these illustrations lifts a suitcase. The weight of the suitcase opposes this motion. This force points down. Since the force supplied by the woman is greater than the weight, there is a net force up, which causes the suitcase to accelerate upward.

equation 1



Newton’s second law

$\boxed{\Sigma\mathbf{F} = m\mathbf{a}}$

$\Sigma\mathbf{F}$ = net force

m = mass

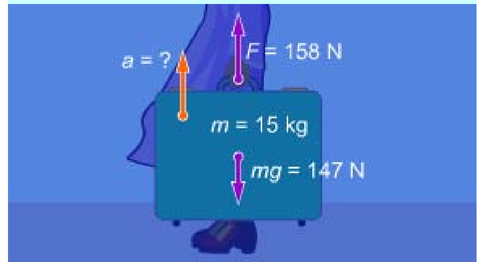
$\mathbf{a}$ = acceleration

Units of force: newtons (N, kg·m/s^2)

In Example 1, the woman lifts the suitcase with a force of 158 N upward. The weight of the suitcase opposes the motion with a downward force of 147 newtons. The two forces act along a line, so we use the convention that up is positive and down is negative, and subtract to find the net force. (If both forces were not acting along a line, you would have to use trigonometry to calculate their components.)

The net force is 11 N, upward. The mass of the suitcase is 15 kg. Newton’s second law can be used to determine the acceleration: It equals the net force divided by the mass. The suitcase accelerates at 0.73 m/s^2 in the direction of the net force, upward.

example 1



What is the suitcase's acceleration?

$\Sigma\mathbf{F} = m\mathbf{a}$

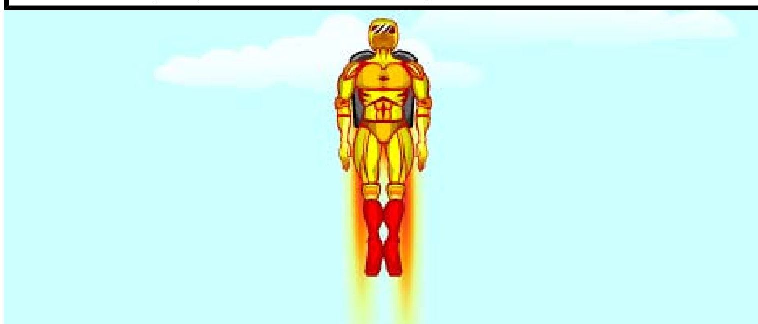
$F + (-mg) = ma$

$a = (F - mg)/m$

$a = (158\text{ N} - 147\text{ N})/(15\text{ kg})$

$a = 0.73\text{ m/s}^2$ (upward)

5.6 - Sample problem: Rocket Guy



Rocket Guy weighs 905 N and his jet pack provides 1250 N of thrust, straight up. What is his acceleration?

Above you see "Rocket Guy," a superhero who wears a jet pack. The jet pack provides an upward force on him, while Rocket Guy's weight points downward.

Variables

All the forces on Rocket Guy are directed along the y axis.

thrust	$F_T = 1250\text{ N}$
weight	$-mg = -905\text{ N}$
mass	m
acceleration	a

What is the strategy?

1. Determine the net force on Rocket Guy.
2. Determine Rocket Guy's mass.
3. Use Newton's second law to find his acceleration.

Physics principles and equations

Newton's second law

$$\Sigma\mathbf{F} = m\mathbf{a}$$

Step-by-step solution

We start by determining the net force on Rocket Guy.

Step	Reason
1. $\Sigma \mathbf{F} = \mathbf{F}_T + m\mathbf{g}$	calculate net vertical force
2. $\Sigma F = F_T + (-mg)$	apply sign conventions
3. $\Sigma F = 1250\text{ N} + (-905\text{ N})$ $\Sigma F = 345\text{ N}$	enter values and add

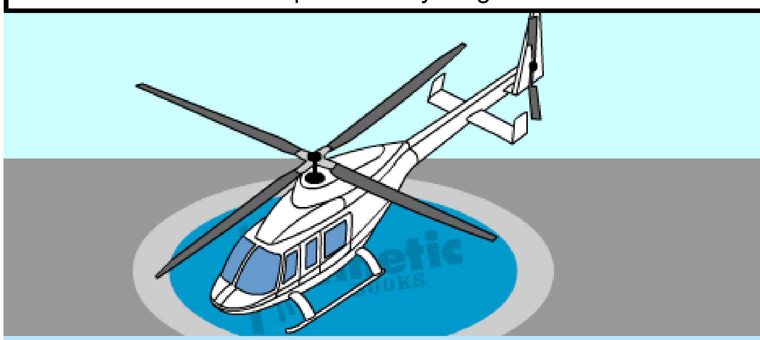
Now we find Rocket Guy's mass.

Step	Reason
4. $m = weight / g$	definition of weight
5. $m = (905\text{ N}) / (9.80\text{ m/s}^2)$ $m = 92.3\text{ kg}$	calculate m

Finally we use Newton's second law to calculate Rocket Guy's acceleration.

Step	Reason
6. $\Sigma \mathbf{F} = m\mathbf{a}$	Newton's second law
7. $a = \Sigma F/m$	solve for a
8. $a = (345\text{ N})/(92.3\text{ kg})$ $a = 3.74\text{ m/s}^2$ (upward)	enter values from steps 3 and 5, and divide

5.7 - Interactive checkpoint: heavy cargo

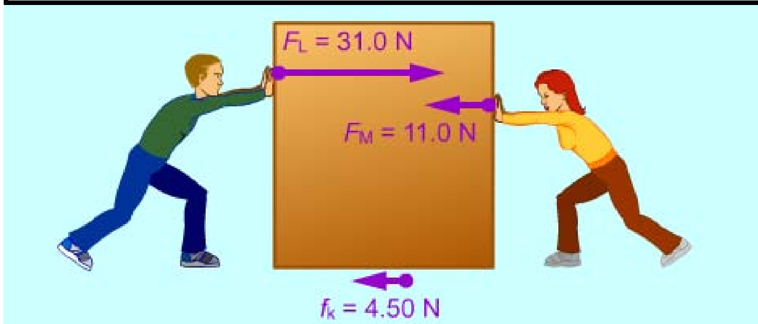


A helicopter of mass 3770 kg can create an upward lift force F. When empty, it can accelerate straight upward at a maximum of 1.37 m/s^2. A careless crewman overloads the helicopter so that it is just unable to lift off. What is the mass of the cargo?

Answer:

$m_c =$ ______ kg

5.8 - Interactive checkpoint: pushing a box



Len pushes toward the right on a 12.0 kg box with a force of magnitude 31.0 N. Martina applies a 11.0 N force on the box in the opposite direction. The magnitude of the kinetic friction force between the box and the very smooth floor is 4.50 N as the box slides toward the right. What is the box's acceleration?

Answer:

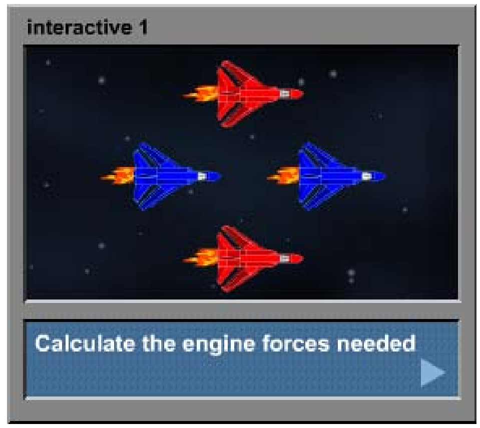
a = ________ m/s^2

5.9 - Interactive problem: flying in formation

The simulation on the right will give you some practice with Newton's second law. Initially, all the space ships have the same velocity. Their pilots want all the ships to accelerate at 5.15 m/s^2. The red ships have a mass of 1.27×10^4 kg, and the blue ships, a mass of 1.47×10^4 kg. You need to set the amount of force supplied by the ships' engines so that they accelerate equally. The masses of the ships do not change significantly as they burn fuel.

Apply Newton's second law to calculate the engine forces needed. The simulation uses scientific notation; you need to enter three-digit leading values. Enter your values and press GO to start the simulation. If all the ships accelerate at 5.15 m/s^2, you have succeeded. Press RESET to try again.

If you have difficulty solving this problem, review Newton's second law.



5.10 - Newton's third law

concept 1

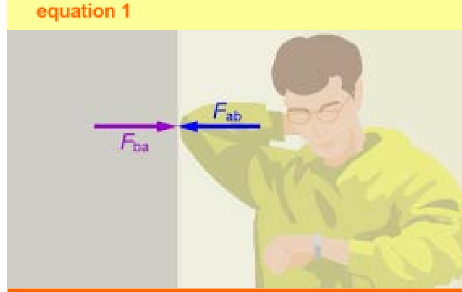
Newton's third law
Forces come in pairs
Equal in strength, opposite in direction
The forces act on different objects

Newton's third law: "To any action there is always an opposite and equal reaction; in other words, the actions of two bodies upon each other are always equal and always opposite in direction."

Newton's third law states that forces come in pairs and that those forces are equal in magnitude and opposite in direction. When one object exerts a force on another, the second object exerts a force equal in magnitude but opposite in direction on the first.

For instance, if you push a button, it pushes back on you with the same amount of force. When someone leans on a wall, it pushes back, as shown in the illustration above.

equation 1



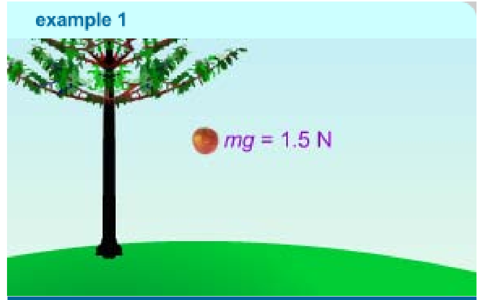
Newton's third law

$$\mathbf{F}_{ab} = -\mathbf{F}_{ba}$$

To illustrate this concept, we use an example often associated with Newton, the falling apple shown in Example 1. The Earth's gravitational force pulls an apple toward the ground and the apple pulls upward on the Earth with an equally strong gravitational force. These pairs of forces are called *action-reaction* pairs, and Newton's third law is often called the action-reaction law.

Force of a on b = opposite of force of b on a

If the forces on the apple and the Earth are equal in strength, do they cause them to accelerate at the same rate? Newton's second law enables you to answer this question. First, objects accelerate due to a net force, and the force of the apple on the Earth is minor compared to other forces, such as those of the Moon or Sun. But, even if the apple were exerting the sole force on the Earth, its acceleration would be very, very small because of the Earth's great mass. The forces are equal, but the acceleration for each body is inversely proportional to its mass.



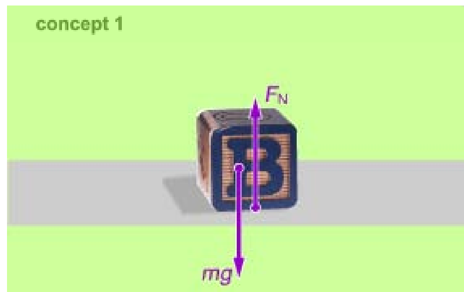
The weight of the apple is 1.5 N. What force does the apple exert on the Earth?
1.5 N upward

5.11 - Normal force

Normal force: When two objects are in direct contact, the force one object exerts in response to the force exerted by the other. This force is perpendicular to the objects' contact surface.

The normal force is a force exerted by one object in direct contact with another. The normal force is a *response force*, one that appears in response to another force. The direction of the force is perpendicular to the surfaces in contact. (One meaning of "normal" is perpendicular.)

A normal force is often a response to a gravitational force, as is the case with the block shown in Concept 1 to the right. The table supports the block by exerting a normal force upward on it. The normal force is equal in magnitude to the block's weight but opposite in direction. The normal force is perpendicular to the surface between the block and the table.

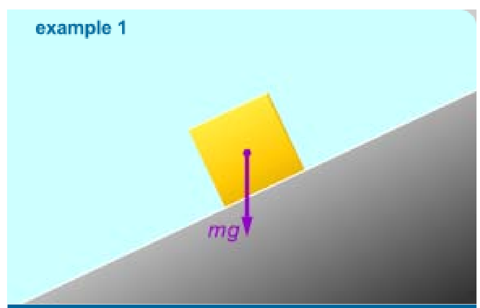


Normal force
Occurs with two objects in direct contact
Perpendicular to surface of contact
Normal force opposes force

You experience the normal force as well. The force of gravity pulls you down, and the normal force of the Earth pushes in the opposite direction. The normal force prevents you from being pulled to the center of the Earth.

Let's consider the direction and the amount of the normal force when you are standing in your classroom. It is equal in magnitude to the force of gravity on you (your weight) and points in the opposite direction. If the normal force were greater than your weight, the net force would accelerate you upward (a surprising result), and if it were less, you would accelerate toward the center of the Earth (equally surprising and likely more distressing). The two forces are equal in strength and oppositely directed, so the amount of the normal force is the same as the magnitude of your weight.

What is the source of the normal force? The weight of the block causes a slight deformation in the table, akin to you lying on a mattress and causing the springs to compress and push back. With a normal force, the deformation occurs at the atomic level as atoms and molecules attempt to "spring back."

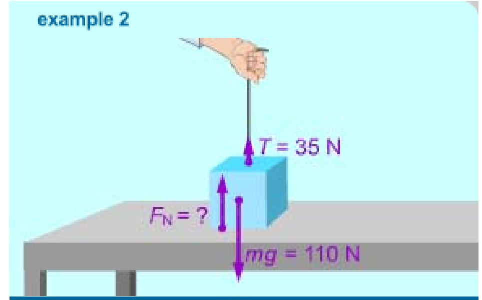


What is the direction of the normal force?
Perpendicular to the surface of the ramp

Normal forces do not just oppose gravity, and they do not have to be directed upward. A normal force is always perpendicular to the surface where the objects are in contact. When you lean against a wall, the wall applies a normal force on you. In this case, the normal force opposes your push and is acting horizontally.

We have discussed normal forces that are acting solely vertically or horizontally. The normal force can also act at an angle, as shown with the block on a ramp in Example 1. The normal force opposes a component of the block's weight, not the full weight. Why? Because the normal force is always perpendicular to the contact surface. The normal force opposes the component of the weight perpendicular to the surface of the ramp.

Example 2 makes a similar point. Here again the normal force and weight are not equal in magnitude. The string pulls up on the block, but not enough to lift it off the surface. Since this reduces the force the block exerts on the table, the amount of the normal force is correspondingly reduced. The force of the string reduces the net downward force on the table to 75 N, so the amount of the normal force is 75 N, as well. The direction of the normal force is upward.



The string supplies an upward force on the block which is resting on the table. What is the normal force of the table on the block?

$\Sigma\mathbf{F} = m\mathbf{a} = \mathbf{0}$

$F_N + T + (-mg) = 0$

$F_N + 35\text{ N} - 110\text{ N} = 0$

$F_N = 75\text{ N (upward)}$

5.12 - Tension

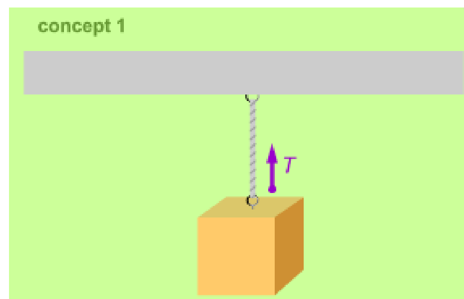
Tension: Force exerted by a string, cord, twine, rope, chain, cable, etc.

In physics textbooks, tension means the pulling force conveyed by a string, rope, chain, tow-bar, or other form of connection. In this section, we will use a rope to illustrate the concept of tension.

The rope in Concept 1 is shown exerting a force on the block; that force is called tension. This definition differs slightly from the everyday use of the word tension, which often refers to forces within a material or object − or a human brain before exams.

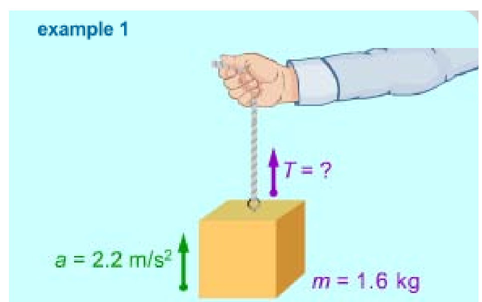
In physics problems, two assumptions are usually made about the nature of tension. First, the force is transmitted unchanged by the rope. The rope does not stretch or otherwise diminish the force. Second, the rope is treated as having no mass (it is massless). This means that when calculating the acceleration of a system, the mass of the rope can be ignored.

Example 1 shows how tension forces can be calculated using Newton’s second law. There are two forces acting on the block: its weight and the tension. The vector sum of those forces, the net force, equals the product of its mass and acceleration. Since the mass and acceleration are stated, the problem solution shows how the tension can be determined.



Tension

Force through rope, string, etc.



What is the amount of tension in the rope?

$\Sigma\mathbf{F} = m\mathbf{a}$

$T + (-mg) = ma$

$T - (1.6\text{ kg})(9.8\text{ m/s}^2) = (1.6\text{ kg})(2.2\text{ m/s}^2)$

$T = 19\text{ N (upward)}$

5.13 - Newton's second and third laws

It might seem that Newton’s third law could lead to the conclusion that forces do **not** cause acceleration, because for every force there is an equal but opposite force. If for every force there is an equal but opposite force, how can there be a net non-zero force? The answer lies in the fact that the forces do not act on the same object. The pair of forces in an action-reaction pair acts on **different** objects. In this section, we illustrate this often confusing concept with an example.

Consider the box attached to the rope in Concept 1. We show two pairs of action-reaction forces. Normally, we draw all forces in the same color, but in this illustration, we draw each pair in a different color. One pair is caused by the force of gravity. The force of the Earth pulls the box down. In turn, the box exerts an upward gravitational force of equal strength on the Earth.

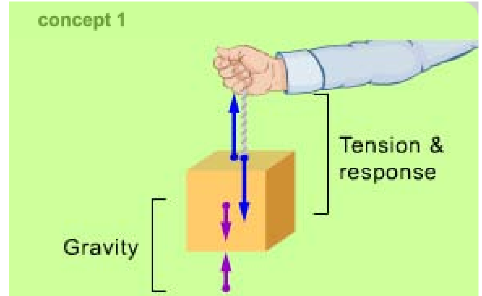
There is also a pair of forces associated with the rope. The tension of the rope pulls up on the box. In response, the box pulls down on the rope. These forces are equal but opposite and form a second action-reaction pair. (Here we only focus on pairs that include forces acting on the box or caused by the box. We ignore other action-reaction pairs present in this example, such as the hand pulling on the rope, and the rope pulling on the hand.)

Now consider only the forces acting **on** the box. This means we no longer consider the forces the box exerts on the Earth and on the rope. The two forces on the box are gravity pulling it down and tension pulling it up. In this example, we have chosen to make the force of tension greater than the weight of the box.

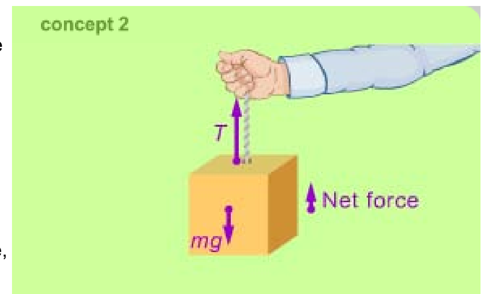
The Concept 2 illustration reflects this scenario: The tension vector is longer than the weight vector, and the resulting net force is a vector upward. Because there is a net upward force on the box, it accelerates in that direction.

Now we will clear up another possible misconception: that the weight of an object resting on a surface and the resulting normal force are an action-reaction pair. They are **not**. Since they are often equal but opposite, they are easily confused with an action-reaction pair. Consider a block resting on a table. The action-reaction pair is the Earth pulling the block down and the block pulling the Earth up. It is not the weight of the block and the normal force.

Here is one way to confirm this: Imagine the block is attached to a rope pulling it up so that it just touches the table. The normal force is now near zero, yet the block's weight is unchanged. If the weight and the normal force are supposed to be equal but opposite, how could the normal force all but disappear? The answer is that the action-reaction pair in question is what is stated above: the equal and opposite forces of gravity between the Earth and the block.



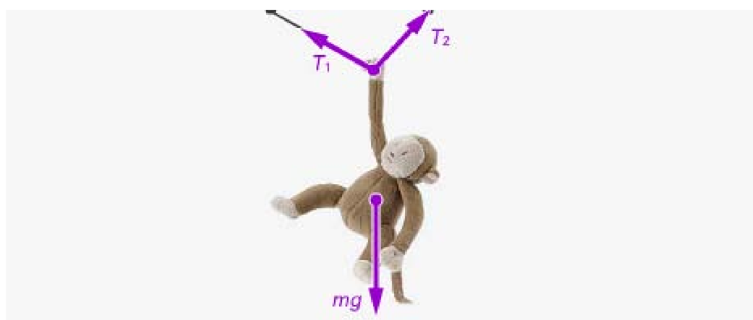
Action-reaction pairs
Two pairs involving box:
· Gravity
· Tension & response



Net force on box
Tension minus weight
Causes acceleration

5.14 - Free-body diagrams

Free-body diagram: A drawing of the external forces exerted on an object.



A monkey hanging from two ropes.

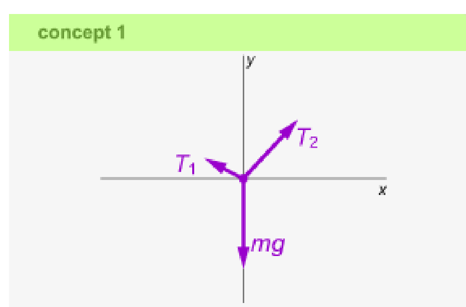
Free-body diagrams are used to display multiple forces acting on an object. In the drawing above, the free body is a monkey, and the free-body diagram in Concept 1 shows the forces acting upon the monkey: the tension forces of the two ropes and the force of gravity.

The diagram only shows the external forces acting on the monkey. There are other forces present in this configuration, such as forces within the monkey, and forces that the monkey exerts. Those forces are not shown; a free-body diagram shows just the forces that act **on** a single object like the monkey.

Although we often draw force vectors where they are applied to an object, in free-body diagrams it is useful to draw the vectors starting from a single point, typically the origin. This allows the components of the vectors to be more easily analyzed. You see this in Concept 1.

Free-body diagrams are useful in a variety of ways. They can be used to determine the magnitudes of forces. For instance, if the mass of the monkey and the orientations of the ropes are known, the tension in each rope can be determined.

When forces act along multiple dimensions, the forces and the resulting acceleration need to be considered independently in each dimension. In the illustration, the monkey is stationary, hanging from two ropes. Since there is no vertical acceleration, there is no **net** force in the vertical dimension. This means the downward force of gravity on the monkey must equal the upward pull of the ropes.

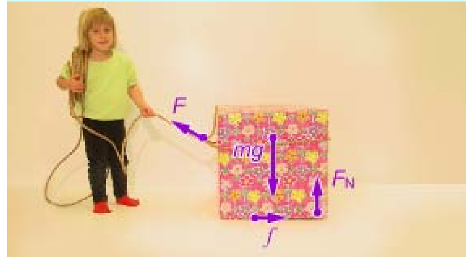


Free-body diagrams
Shows all external forces acting on body
Often drawn from the origin

The two ropes also pull horizontally (along the x axis). Because the monkey is not accelerating horizontally, these horizontal forces must balance as well. By considering the forces acting in both the horizontal and vertical directions, the tensions of the ropes can be determined.

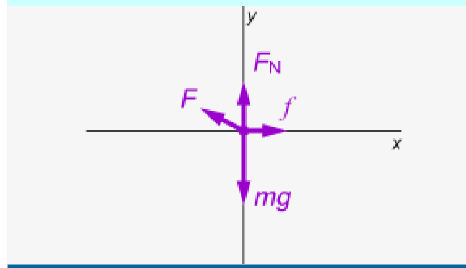
In Example 1, one of the forces shown is friction, f. Friction acts to oppose motion when two objects are in contact.

example 1



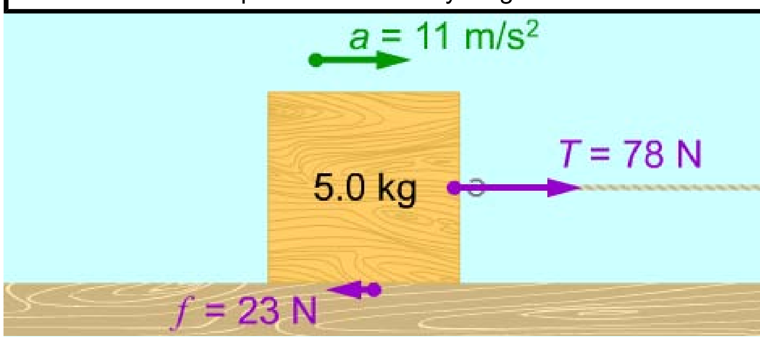
Draw a free-body diagram of the forces on the box.

example 2



Free-body diagram of forces on box

5.15 - Interactive problem: free-body diagram



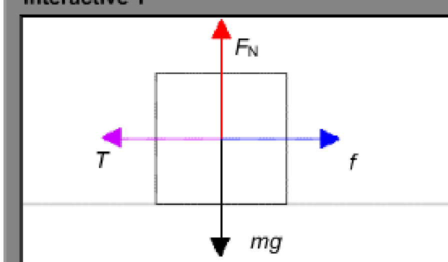
A rope pulls the block against friction. Draw a free-body diagram. The block accelerates at 11 m/s^2 if the diagram is correct.

In this section, you practice drawing a free-body diagram. Above, you see the situation: A block is being pulled horizontally by a rope. It accelerates to the right at 11 m/s^2. In the simulation on the right, the force vectors on the block are drawn, but each one points in the wrong direction, has the wrong magnitude, or both. We ignore the force of air resistance in this simulation.

Your job is to fix the force vectors. You do this by clicking on the heads of the vectors and dragging them to point in the correct direction. (To simplify your work, they "snap" to vertical and horizontal orientations, but you do need to drag them close before they will snap.) You change both their lengths (which determine their magnitudes) and their directions with the mouse.

The mass of the block is 5.0 kg. The tension force T is 78 N and the force of friction f is 23 N. The friction force acts opposite to the direction of the motion. Calculate the magnitudes of the weight mg and the normal force F_N to the nearest newton, and then drag the heads of the vectors to the correct positions, or click on the up and down arrow buttons, and press GO. If you are correct, the block will accelerate to the right at 11 m/s^2. If not, the block will move based on the net force as determined by your vectors as well as its mass. Press RESET to try again.

interactive 1

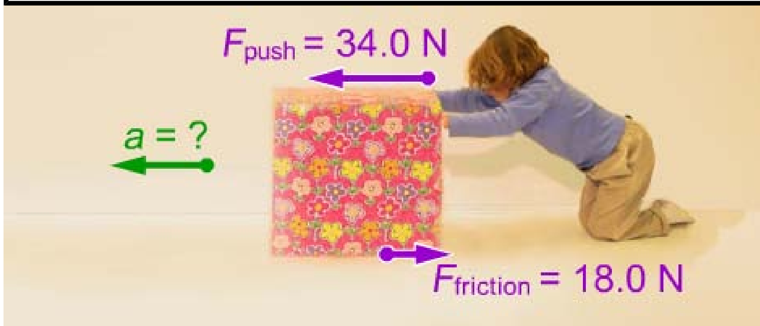


Free-body diagram
Drag vectors to achieve an acceleration

There is more than one way to arrange the vectors to create the same acceleration, but there is only one arrangement that agrees with all the information given.

If you have difficulty solving this problem, review the sections on weight and normal force, and the section on free-body diagrams.

5.16 - Sample problem: pushing a box horizontally

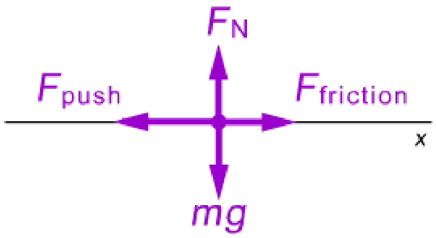


What is the acceleration of the 15.0 kg box?

The girl shown above is pushing the box horizontally, causing it to accelerate. The force of friction opposes this movement. You are asked to find the acceleration of the box.

With problems like this, we start with a free-body diagram.

Draw a free-body diagram



The free-body diagram shows all the forces acting on the box. In solving the problem, we use only the horizontal forces from the push and friction in our calculations. There is no net vertical force because the downward force of weight is balanced by the upward normal force. The fact that there is no vertical acceleration confirms that there is no net vertical force.

Variables

Since the motion is in one dimension, we use signs to indicate direction. As usual, we take the positive direction to be to the right.

force of push	$-F_{push} = -34.0$ N
force of friction	$F_{friction} = 18.0$ N
mass	$m = 15.0$ kg
acceleration	a

What is the strategy?

1. Draw a free-body diagram.
2. Calculate the net force on the box.
3. Use Newton's second law to calculate the acceleration of the box.

Physics principles and equations

Newton's second law

$$\Sigma \mathbf{F} = m\mathbf{a}$$

Step-by-step solution

As noted, we use the convention that forces to the right are positive and those to the left are negative. A more rigorous approach would be to calculate the vector components of these forces using the cosine of $0°$ for the frictional force and the cosine of $180°$ for the pushing force. The result would be x components of 18.0 N and −34.0 N (the same conclusion we reached via inspection and convention). Many instructors prefer this approach. It does not change the answer to the problem, but the component method is more rigorous, and is required to solve more difficult problems.

Step	Reason
1. $\Sigma\mathbf{F} = \mathbf{F}_{push} + \mathbf{F}_{friction}$	net horizontal force
2. $\Sigma F = –34.0\text{ N} + 18.0\text{ N}$ $\Sigma F = –16.0\text{ N}$	enter values and add
3. $\Sigma F = ma$	Newton's second law
4. $a = \Sigma F/m$	solve for a
5. $a = (–16.0\text{ N}) / (15.0\text{ kg})$	enter values
6. $a = –1.07\text{ m/s}^2$	division

5.17 - Interactive problem: lifting crates

The helicopter on the right is being used as a scale, making it one of the more expensive scales in the world, we suspect. This simulation includes three crates; each has a slightly different mass. Your assignment is to find the crate with a mass of 661 kg. Do this by lifting each crate with the helicopter and noting the acceleration. The helicopter lifts each crate with a force of 10,748 N via the tension in the cable. The resulting acceleration of each crate will let you calculate its mass.

Click on the graphic to start the simulation. To determine the answer, drag the helicopter to each of the three crates and press GO to make the helicopter lift the crate. Record the acceleration of each crate and use the acceleration to calculate the mass. When you have found the crate with a mass of 661 kg, select it by clicking on it. The simulation will tell you whether you clicked on the correct one.

If you cannot solve the problem, review Newton's second law and the section on weight.

interactive 1

Lift the crates to find a mass of 661 kg

5.18 - Friction

Friction: A force that resists the motion of one object sliding past another.

Friction between the buffalo's back and the tree scratches an itch.

If you push a cardboard box along a wooden floor, you have to push to overcome the force of friction. This force makes it harder for you to slide the box. The force of friction opposes any force that can cause one object to slide past another. There are two types of friction: static and kinetic. These forces are discussed in more depth in other sections. In this section, we discuss some general properties of friction.

The amount of friction depends on the materials in contact. For example, the box would slide more easily over ice than wood. Friction is also proportional to the normal force. For a box on the floor, the greater its weight, the greater the normal force, which increases the force of friction.

Humans expend many resources to combat friction. Motor oil, Teflon™, WD-40™, Tri-Flo™ and many other products are designed to reduce this force. However, friction can be very useful. Without it, a nail would slip out of a board, the tires of a car would not be able to “grip” the road, and you would not be able to walk.

Friction exists even between seemingly smooth surfaces. Although a surface may appear smooth, when magnified sufficiently, any surface will look bumpy or rough, as the illustration in Concept 2 on the right shows. The magnified picture of the “smooth” crystal reveals its microscopic “rough” texture. Friction is a force caused by the interaction of molecules in two surfaces.

You might think you can defeat friction by creating surfaces that are highly polished. Instead, you may get an effect called *cold welding*, in which the two highly polished materials fuse together. Cold welding can be desirable, as when an aluminum connector is crimped onto a copper wire to create a strong electrical connection.

Objects can also move in a fashion that is called *slip and slide*. They slide for a while, stick, and then slide some more. This phenomenon accounts for both the horrid noise generated by fingernails on a chalkboard and the joyous noise of a violin. (Well, joyous when played by some, chalkboard-like when played by others.)

concept 1

Friction
Force that opposes "sliding" motion
Varies by materials in contact
Proportional to normal force

concept 2

Friction
Microscopic properties determine friction force

5.19 - Static friction

Static friction: A force that resists the sliding motion of two objects that are stationary relative to one another.

Imagine you are pushing a box horizontally but cannot move it due to friction. You are experiencing a response force called static friction. If you push harder and harder, the amount of static friction will increase to exactly equal − but not exceed − the amount of horizontal force you are supplying. For the two surfaces in contact, the friction will increase up to some maximum amount. If you push hard enough to exceed the maximum amount of static friction, the box will slide.

For instance, let's say the maximum amount of static friction for a box is 30 newtons. If you push with a force of 10 newtons, the box does not move. The force of static friction points in the opposite direction of your force and is 10 newtons as well. If it were less, the box would slide in the direction you are pushing. If it were greater, the box would accelerate toward you. The box does not move in either direction, so the friction force is 10 newtons. If you push with 20 newtons of force, the force of static friction is 20 newtons, for the same reasons.

You keep pushing until your force is 31 newtons. You have now exceeded the maximum force of static friction and the box accelerates in the direction of the net force. The box will continue to experience friction once it is sliding, but this type of friction is called kinetic friction.

concept 1

Static friction
Force opposing sliding when no motion
Balances "pushing" force until object slides
Maximum static friction proportional to:
· coefficient of static friction
· normal force

Static friction occurs when two objects are motionless relative to one another. Often, we want to calculate the maximum amount of static friction so that we know how much force we will have to apply to get the object to move. The equation in Equation 1 enables you to do so. It depends on two values. One is the normal force, the perpendicular force between the two surfaces. The second is called the *coefficient of static friction*.

Engineers calculate this coefficient empirically. They place an object (say, a car tire) on top of another surface (perhaps ice) and measure how hard they need to push before the object starts to move. Coefficients of friction are specific to the two surfaces. Some examples of coefficients of static friction are shown in the table in Equation 2.

You might have noticed a fairly surprising fact: The amount of surface area between the two objects does not enter into the calculation of maximum static friction. In principle, whether a box of a given mass has a surface area of one square centimeter or one square kilometer, the maximum amount of static friction is constant. Why? With the greater contact area, the normal and frictional forces per unit area diminish proportionally.

equation 1

Static friction

$$f_{s,max} = \mu_s F_N$$

$f_{s,max}$ = maximum static friction
μ_s = coefficient of static friction
F_N = normal force

equation 2

Coefficient of static friction	
Tires on dry pavement	0.90
Tires on wet pavement	0.42
Glass on glass	0.94
Steel on steel	0.78
Oak on oak	0.54
Waxed ski on dry snow	0.04
Teflon™ on Teflon™	0.04

Coefficients of static friction

example 1

Anna is pushing but the box does not move. What is the force of static friction?

f_s = 7 N to the right

example 2

What is the maximum static

friction force? The coefficient of static friction for these materials is 0.31.

$f_{s,max} = \mu_s F_N$

$f_{s,max} = (0.31)(27\ N)$

$f_{s,max} = 8.4\ N$

5.20 - Kinetic friction

Kinetic friction: Friction when an object slides along another.

Kinetic friction occurs when two objects slide past each other. The magnitude of kinetic friction is less than the maximum amount of static friction for the same objects. Some values for coefficients of kinetic friction are shown in Equation 2 to the right. These are calculated empirically and do not vary greatly over a reasonable range of velocities.

Like static friction, kinetic friction always opposes the direction of motion. It has a constant value, the product of the normal force and the coefficient of kinetic friction.

In Example 1, we state the normal force. Note that the normal force in this case does **not** equal the weight; instead, it equals a component of the weight. The other component of the weight is pulling the block down the plane.

concept 1

Kinetic friction

Friction opposing sliding in motion
Force constant as object slides
Proportional to:
· coefficient of kinetic friction

· normal force

equation 1

Kinetic friction

$$f_k = \mu_k F_N$$

f_k = force of kinetic friction

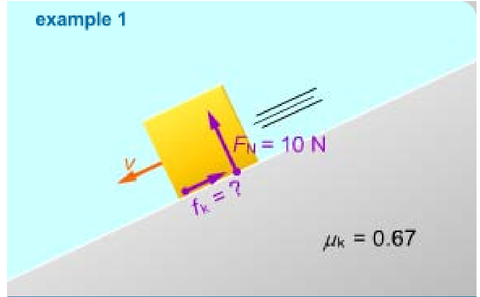
μ_k = coefficient of kinetic friction

F_N = normal force

equation 2

Coefficient of kinetic friction	
Tires on dry pavement	0.85
Tires on wet pavement	0.36
Glass on glass	0.40
Steel on steel	0.42
Oak on oak	0.32
Waxed ski on dry snow	0.03
Teflon® on Teflon®	0.04

Coefficients of kinetic friction



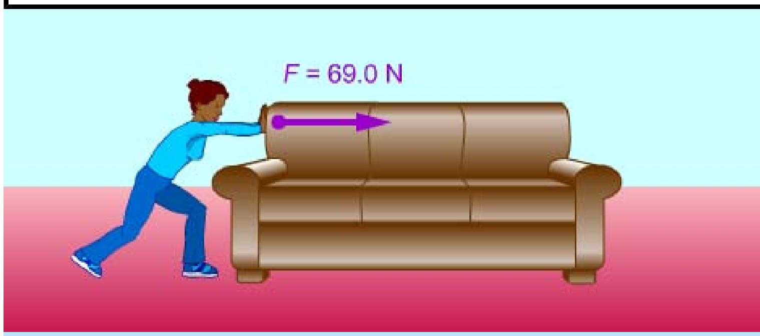
What is the force of friction?

$f_k = \mu_k F_N$

f_k = (0.67)(10 N)

f_k = 6.7 N (pointing up the ramp)

5.21 - Interactive checkpoint: moving the couch

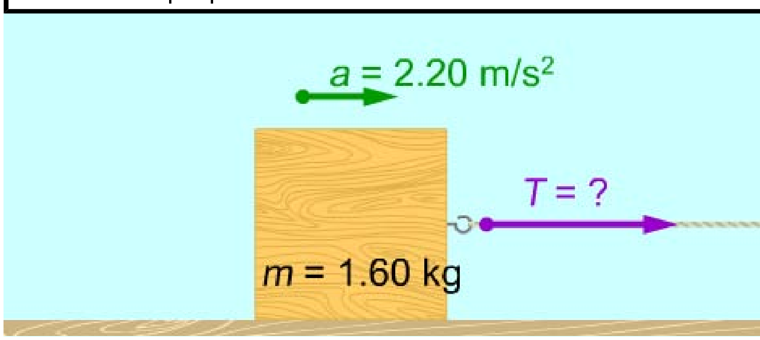


While rearranging your living room, you push your couch across the floor at a constant speed with a horizontal force of 69.0 N. You are using special pads on the couch legs that help it slide easier. If the couch has a mass of 59.5 kg, what is the coefficient of kinetic friction between the pads and the floor?

Answer:

μ_k =

5.22 - Sample problem: friction and tension

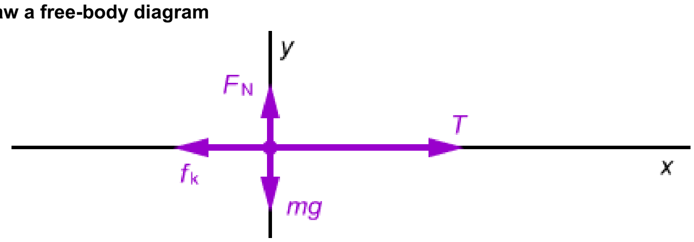


The coefficient of kinetic friction is 0.200. What is the magnitude of the tension force in the rope?

Above, you see a block accelerating to the right due to the tension force applied by a rope. What is the magnitude of tension the rope applies to the block?

Starting this type of problem with a free-body diagram usually proves helpful.

Draw a free-body diagram



Variables

	x component	y component
normal force	0	$F_N = mg$
acceleration	$a = 2.20\ \text{m/s}^2$	0
tension	T	0
friction force	$-f_k$	0
mass	$m = 1.60\ \text{kg}$	
coefficient of kinetic friction	$\mu_k = 0.200$	

What is the strategy?

1. Draw a free-body diagram.
2. Find an expression for the net force on the block.
3. Substitute the net force into Newton's second law to find the tension.

Are there any useful relationships?

Since the surface is horizontal, the amount of normal force equals the weight of the block.

Physics principles and equations

Newton's second law

$$\Sigma\mathbf{F} = m\mathbf{a}$$

The magnitude of the force of kinetic friction is found by

$$f_k = \mu_k F_N$$

Step-by-step solution

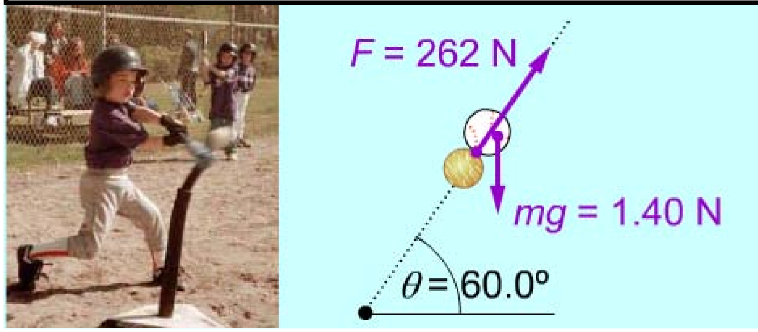
We begin by determining the net horizontal force on the block.

Step	Reason
1. $\Sigma F = T + (-f_k)$	net horizontal force
2. $f_k = \mu_k F_N$	equation for kinetic friction
3. $\Sigma F = T - \mu_k F_N$	substitute equation 2 into 1
4. $\Sigma F = T - \mu_k mg$	enter value of F_N

Now we substitute the net force just found into Newton's second law. This allows us to solve for the tension force.

Step	Reason
5. $\Sigma F = ma$	Newton's second law
6. $T - \mu_k mg = ma$	substitute equation 4 into 5
7. $T = \mu_k mg + ma$	solve for tension
8. $T = (0.200)(1.60\ \text{kg})(9.80\ \text{m/s}^2) + (1.60\ \text{kg})(2.20\ \text{m/s}^2)$	enter values
9. $T = 6.66\ \text{N}$	evaluate

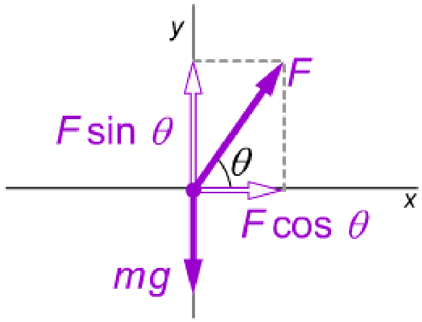
5.23 - Sample problem: a force at an angle



What is the magnitude of the net force on the ball along each axis, and what is the ball's acceleration along each axis?

Above, you see a bat hitting a ball at an angle. You are asked to find the net force and the acceleration of the ball along the x and y axes.

Draw a free-body diagram



The forces on the ball are its weight down and the force of the bat at the angle θ to the x axis.

Variables

	x component	y component
weight	0	$mg \sin 270° = -1.40$ N
force	$F \cos \theta$	$F \sin \theta$
acceleration	a_x	a_y
force	$F = 262$ N	
angle	$\theta = 60.0°$	
mass	$m = mg/g = (1.40\text{ N}) / (9.80\text{ m/s}^2) = 0.143$ kg	

What is the strategy?

1. Draw a free-body diagram.
2. Use trigonometry to calculate the net force on the ball along each axis.
3. Use Newton's second law to find the acceleration of the ball along each axis. The mass of the ball is not given, but you can determine it because you are told its weight. We do this in the variables table.

Physics principles and equations

Newton's second law

$$\Sigma\mathbf{F} = m\mathbf{a}$$

Step-by-step solution

We begin by calculating the net force along the x axis.

Step	Reason
1. $\Sigma F_x = F \cos \theta$	net force along x axis
2. $\Sigma F_x = (262\text{ N})(\cos 60.0°)$	x component of force
3. $\Sigma F_x = 131$ N	evaluate

We next calculate the force along the *y* axis. In this case, there are two forces to consider.

Step	Reason
4. $\Sigma F_y = F \sin\theta + (-1.40\text{ N})$	net force along *y* axis
5. $\Sigma F_y = (262\text{ N})(\sin 60.0°) + (-1.40\text{ N})$	enter values
6. $\Sigma F_y = 225\text{ N}$	evaluate

Now we calculate the acceleration along the *x* axis, using Newton's second law.

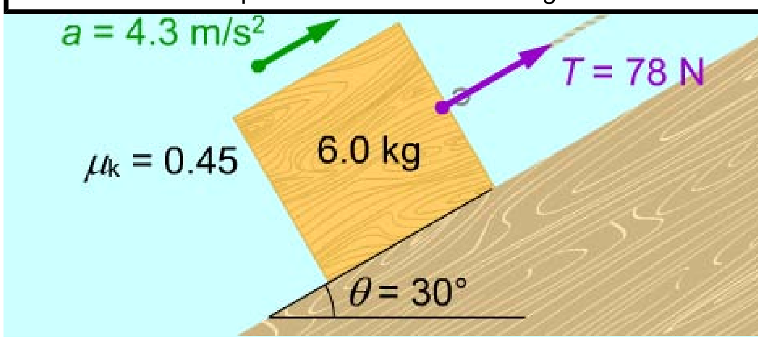
Step	Reason
7. $\Sigma F_x = ma_x$	Newton's second law
8. $a_x = \Sigma F_x/m$	solve for a_x
9. $a_x = (131\text{ N}) / (0.143\text{ kg})$	enter values from step 3 and table
10. $a_x = 916\text{ m/s}^2$	division

We calculate the acceleration along the *y* axis.

Step	Reason
11. $\Sigma F_y = ma_y$	Newton's second law
12. $a_y = \Sigma F_y/m$	solve for a_y
13. $a_y = (225\text{ N})/(0.143\text{ kg})$	enter values from step 6, table
14. $a_y = 1570\text{ m/s}^2$	division

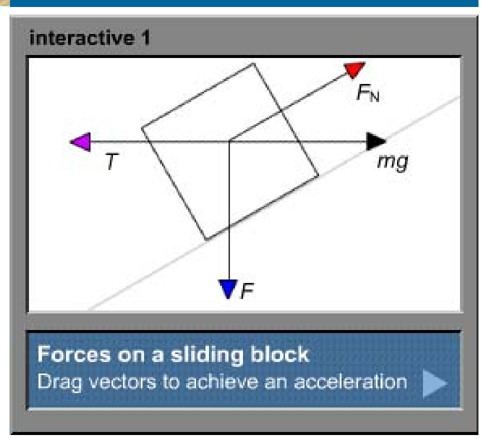
The acceleration values may seem very large, but this is the acceleration during the brief moment the bat is in contact with the ball.

5.24 - Interactive problem: forces on a sliding block



A rope pulls the block up the ramp. Draw a free-body diagram of the forces on the block. If the diagram is correct, the block will accelerate up the ramp at 4.3 m/s^2.

Above, you see an illustration of a block that is being pulled up a ramp by a rope. In the simulation on the right, the force vectors on the block are drawn, but they are in the wrong directions, have the wrong magnitudes, or both. Your job is to fix the force vectors. If you do this correctly, the block will accelerate up the ramp at a rate of 4.3 m/s^2. If not, the block will move due to the net force as determined by your vectors as well as its mass.

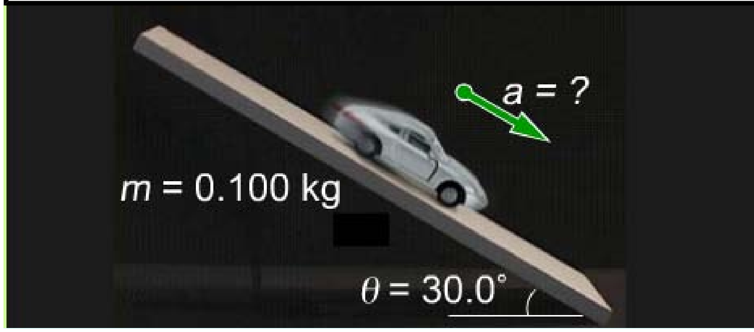


The mass of the block is 6.0 kg. The amount of tension from the rope is 78 N and the coefficient of kinetic friction is 0.45. The angle the ramp makes with the horizontal is 30°. Calculate (to the nearest newton) the directions and magnitudes of the weight, normal force and friction force. Drag the head of a vector to set its magnitude and direction. You can also set the magnitudes in the control panel. The vectors will "snap" to angles.

When you have arranged all the vectors, press the GO button. If your free-body diagram is accurate, the block will accelerate up the ramp at 4.3 m/s^2. Press RESET to try again.

There is more than one way to set the vectors to produce the same acceleration, but only one arrangement agrees with all the information given. If you have difficulty solving this problem, review the sections on kinetic friction and the normal force, and the sample problem involving a force at an angle.

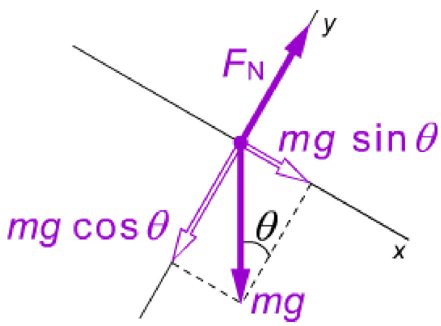
5.25 - Sample problem: moving down a frictionless plane

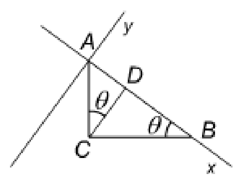


What is the car's acceleration?

Above, you see a toy car going down an inclined plane. The diagram shows the mass of the car and the angle the plane makes with the horizontal. You are asked to calculate the car's acceleration. In this problem, ignore any friction or air resistance, as well as any energy consumed by the rotation of the wheels.

Draw a free-body diagram





By rotating the axes so that the x axis is parallel to the car's motion down the ramp, we make the forces along the y axis sum to zero. (These two forces are the y component of the car's weight and the normal force from the ramp.) Rotating the axes means there is a net force only along the x axis, and this reduces the steps required to solve the problem.

It may be a little difficult to see why θ, the angle that the plane makes with the horizontal, is the same as the angle θ in the free-body diagram. The drawing to the right of the free-body diagram uses two similar right triangles to show why this is true. The triangle ABC has one leg (AC) that is the weight vector, and its hypotenuse (AB) lies along the x axis. The hypotenuse of the smaller triangle ACD is the weight vector. These are both right triangles and share a common angle at A, so they are similar.

It is often useful to check this angle with the situation shown. At a $30°$ angle, the y component of the weight is larger than the x component (the cosine of $30°$ is greater than the sine of $30°$). Looking at the picture above, this is what you would expect. The component of the weight down the plane is less than the component on the plane. It may help to push it to the extreme: What would you expect at a $0°$ angle? At $90°$?

Variables

With the axes rotated and θ as shown, the x component of the weight is computed using the sine, and the y component with the cosine. (Without the rotation, the x component would be calculated with the cosine, and the y component with the sine.)

	x component	y component
weight	$mg \sin \theta$	$-mg \cos \theta$
normal force	0 N	F_N
acceleration	a	0 m/s^2
mass	$m = 0.100$ kg	
angle	$\theta = 30.0°$	

What is the strategy?

1. Draw a free-body diagram, rotating the axes so the x axis is parallel to the motion of the car.
2. Use trigonometry to calculate the net force on the car.
3. Use Newton's second law to determine the acceleration of the car.

Physics principles and equations

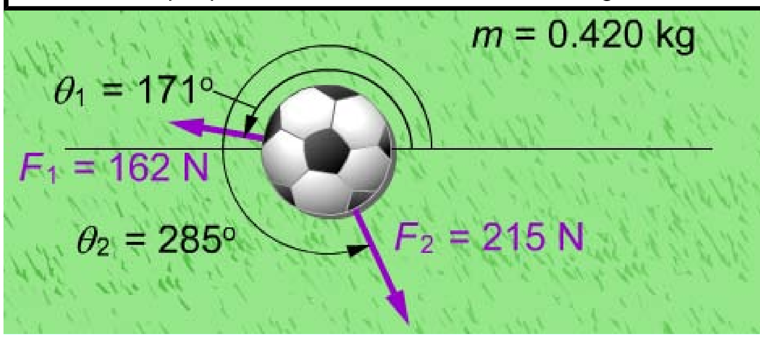
Newton's second law

$$\Sigma \mathbf{F} = m\mathbf{a}$$

Step-by-step solution

Step	Reason
1. $\Sigma F_x = mg \sin \theta$	net force along x axis
2. $\Sigma F_x = (0.100\text{ kg})(9.80\text{ m/s}^2)(\sin 30.0°)$	enter values
3. $\Sigma F_x = 0.490\text{ N}$	evaluate
4. $\Sigma F_x = ma$	Newton's second law
5. $a = \Sigma F_x / m$	solve for a
6. $a = (0.490\text{ N}) / (0.100\text{ kg})$	enter values
7. $a = 4.90\text{ m/s}^2$ (down the plane)	division

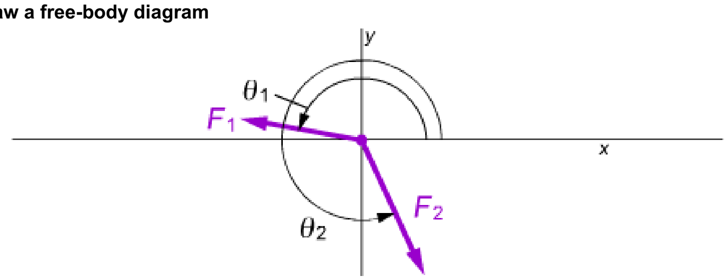
5.26 - Sample problem: two forces at different angles



Two players kick the soccer ball lying on the ground with the forces shown. What are the x and y components of the resulting acceleration of the ball?

You see a top-down view of the ball and the forces exerted on it by the players. You are asked to find the x and y components of the resulting acceleration. When solving the problem, ignore the force of friction.

Draw a free-body diagram



Variables

We use trigonometry to calculate the components of each force. The angles θ_1 and θ_2 are measured from the positive x axis.

	x component	y component
force 1	$F_1 \cos \theta_1$	$F_1 \sin \theta_1$
force 2	$F_2 \cos \theta_2$	$F_2 \sin \theta_2$
acceleration	a_x	a_y
mass	0.420 kg	
force 1 angle	$\theta_1 = 171°$	
force 2 angle	$\theta_2 = 285°$	
force 1	$F_1 = 162\text{ N}$	
force 2	$F_2 = 215\text{ N}$	

What is the strategy?

1. Draw a free-body diagram.
2. Calculate the net force on the ball along each axis by finding the components of the two forces using trigonometry.
3. Use Newton's second law to find the acceleration of the ball along each axis.

Physics principles and equations

Newton's second law

$$\Sigma\mathbf{F} = m\mathbf{a}$$

Step-by-step solution

We begin by calculating the net force along the *x* axis.

Step	Reason
1. $\Sigma F_x = F_1 \cos\theta_1 + F_2 \cos\theta_2$	net force along x axis
2. $\Sigma F_x = 162\cos(171°)\text{N} + 215\cos(285°)\text{N}$	enter values
3. $\Sigma F_x = -104\text{ N}$	evaluate

Next we calculate the net force along the *y* axis.

Step	Reason
4. $\Sigma F_y = F_1 \sin\theta_1 + F_2 \sin\theta_2$	net force along y axis
5. $\Sigma F_y = 162\sin(171°)\text{N} + 215\sin(285°)\text{N}$	enter values
6. $\Sigma F_y = -182\text{ N}$	evaluate

Using Newton's second law and the net force in the *x* dimension, calculated above, we find the acceleration in the *x* dimension.

Step	Reason
7. $\Sigma F_x = ma_x$	Newton's second law
8. $a_x = \Sigma F_x / m$	solve for a_x
9. $a_x = (-104\text{ N}) / (0.420\text{ kg})$	enter values
10. $a_x = -248\text{ m/s}^2$	division

And finally we calculate the acceleration in the *y* dimension.

Step	Reason
11. $\Sigma F_y = ma_y$	Newton's second law
12. $a_y = \Sigma F_y / m$	solve for a_y
13. $a_y = (-182\text{ N}) / (0.420\text{ kg})$	enter values
14. $a_y = -433\text{ m/s}^2$	division

5.27 - Interactive checkpoint: sledding

A child sits on a sled on a frictionless, icy hill that is inclined at 25.0° from the horizontal. The mass of the child and sled is 36.5 kg. What is the magnitude of the normal force of the hill on the child? At what rate does the child accelerate down the hill?

Answer:

F_N = ______ N

a = ______ m/s^2

5.28 - Hooke's law and spring force

You probably already know a few basic things about springs: You stretch them, they pull back on you. You compress them, they push back.

As a physics student, though, you are asked to study springs in a more quantitative way. Let's consider the force of a spring using the configuration shown in Concept 1. Initially, no force is applied to the spring, so it is neither stretched nor compressed. When no force is applied, the end of the spring is at a position called the rest point (sometimes called the equilibrium point).

Then we stretch the spring. In the illustration to the right, the hand pulls to the right, so the end of the spring moves to the right, away from its rest point, and the spring pulls back to the left.

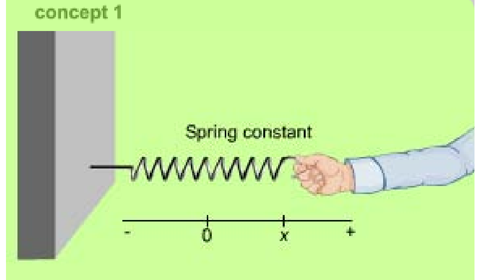
Hooke's law is used to determine how much force the spring exerts. It states that the amount of force is proportional to how far the end of the spring is stretched or compressed away from its rest point. Stretch the end of the spring twice as far from its rest point, and the amount of force is doubled.

The amount of force is also proportional to a spring constant, which depends on the construction of the spring. A "stiff" spring has a greater spring constant than one that is easier to stretch. Stiffer springs can be made from heavier gauge materials. The units for spring constants are newtons per meter (N/m).

The equation for Hooke's law is shown in Equation 1. The spring constant is represented by k. The displacement of the end of the spring is represented by x. At the rest position, $x = 0$. When the spring is stretched, the displacement of the end of the spring has a positive x value. When it is compressed, x is negative.

Hooke's law calculates the magnitude of the spring force. The equation has a negative sign to indicate that the force of a spring is a *restoring force*, which means it acts to restore the end of the spring to its rest point. Stretch a spring and it will pull back toward the rest position; compress a spring, and it will push back toward the rest position. The direction of the force is the opposite of the direction of the displacement.

concept 1



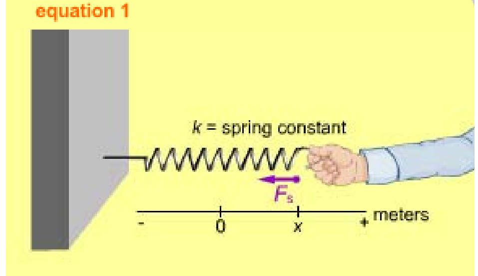
Spring force

Force exerted by spring depends on:

· How much it is stretched or compressed

· Spring constant

equation 1



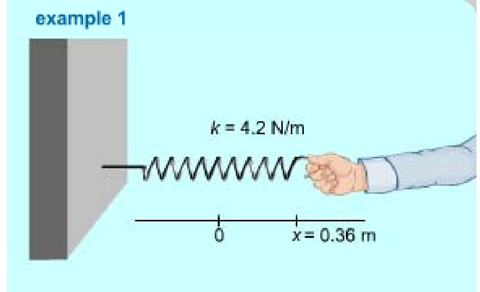
Hooke's law

$$F_s = -kx$$

F_s = spring force

k = spring constant

x = displacement of end from rest point



What is the force exerted by the spring?

$F_s = -kx$

$F_s = -(4.2\ \text{N/m})(0.36\ \text{m})$

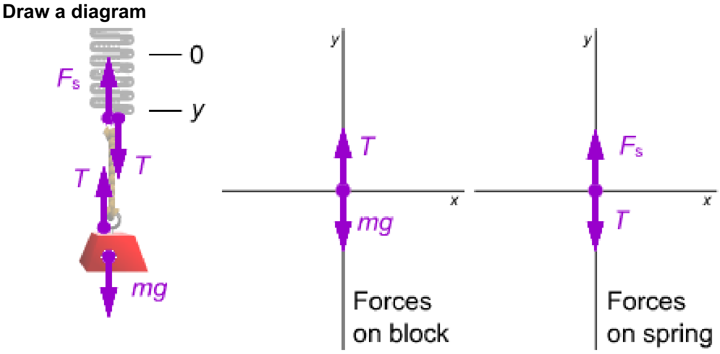
$F_s = -1.5\ \text{N}$ (to the left)

5.29 - Sample problem: spring force and tension

What is the amount of tension in the rope? What is the position of the end of the spring away from its rest position?

You see a block hanging from a rope attached to a spring. The block is stationary. You are asked to determine the tension in the rope and the position of the end of the spring relative to its rest point.

Draw a diagram



Above on the left, we draw the forces: the weight of the block, the tension forces in the rope, and the spring force. The tension forces are equal in magnitude, so we use the same variable T for each of them. Then we draw two free-body diagrams, one for the block and one for the lower end of the spring.

Variables

weight	$-mg = -98.0\ \text{N}$
tension	T
spring force	F_s
spring constant	$k = 535\ \text{N/m}$
displacement	y

What is the strategy?

1. Draw a free-body diagram.
2. Apply Newton's second law to calculate the tension force in the rope.
3. Then apply Newton's second law with Hooke's law to find the position of the end of the spring.

Physics principles and equations

Newton's second law

$$\Sigma \mathbf{F} = m\mathbf{a}$$

Hooke's law, applied along the y axis

$$F_s = -ky$$

The forces sum to zero for each object because there is no acceleration.

Step-by-step solution

We start by determining the magnitude of tension in the rope, using the fact that the block is stationary.

Step	Reason
1. $T + (-mg) = 0$	no acceleration means no net force
2. $T = 98.0$ N	enter value for weight into step 1 to determine tension

Since the acceleration of the end of the spring is also zero, the forces again sum to zero. We then use Hooke's law and solve for the displacement.

Step	Reason
3. $F_s + (-T) = 0$	no net force
4. $-ky + (-T) = 0$	Hooke's law
5. $y = -T / k$	solve for y
6. $y = -(98.0\text{ N}) / (535\text{ N/m})$	enter values
7. $y = -0.183$ m	evaluate

5.30 - Air resistance

Air resistance: A force that opposes motion in air.

If you parachute, or bike or ski, you have experienced air resistance. In each of these activities, you move through a fluid − air − that resists your motion. As you move through the air, you collide with the molecules that make up the atmosphere. Although air is not very dense and the molecules are very small, there are so many of them that their effects add up to a significant force. The sum of all these collisions is the force called air resistance.

Unlike kinetic friction, air resistance is not constant but increases as the speed of the object increases. The force created by air resistance is called *drag.*

The formula in Equation 1 supplies an approximation of the force of air resistance for objects moving at relatively high speeds through air. For instance, it is a relevant equation for the skysurfer shown in Concept 1, or for an airplane. The resistance is proportional to the square of the speed and to the cross sectional area of the moving object. (For the skysurfer, the board would constitute the main part of the cross sectional area.) It is also proportional to an empirically determined constant called the *drag coefficient*.

The shape of an object determines its drag coefficient. A significant change in speed can change the drag coefficient, as well. Aerospace engineers definitely earn their keep by analyzing air resistance using powerful computers. They also use wind tunnels to check their computational results.

Another interesting implication of the drag force equation is that objects will reach what is called *terminal velocity*. Terminal velocity is the maximum speed an object reaches when falling. The drag force increases with speed while the force of gravity is constant; at some point, the upward drag force equals the downward force of gravity. When this occurs, there is no net force and the object ceases to accelerate and maintains a constant speed. The equation for calculating terminal velocity is shown in Equation 2. It is derived by setting the drag force equal to the object's weight and solving for the speed.

concept 1

Air resistance
Drag force opposes motion in air
Force increases as speed increases

concept 2

Terminal velocity
Drag force equals weight

Research has actually determined that cats reach terminal velocity after falling six stories. In fact, they tend to slow down after six stories. Here's why this occurs: The cat achieves terminal velocity and then relaxes a little, which expands its cross sectional area and increases its drag force. As a result, it slows down. One has to admire the cat for relaxing in such a precarious situation (or perhaps doubt its intelligence). If you think this may be an urban legend, consult the *Journal of the American Veterinary Association*, volume 191, page 1399.

equation 1

Air resistance

$$F_D = \tfrac{1}{2}C\rho A v^2$$

F_D = drag force
C = drag coefficient for object
ρ = air density
A = cross-sectional area
v = velocity

equation 2

Terminal velocity

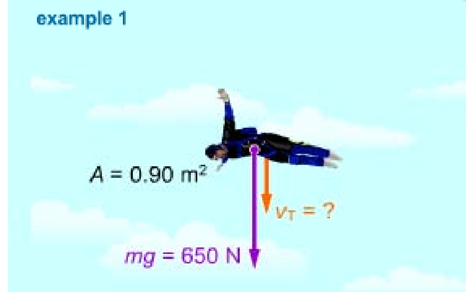
$$v_T = \sqrt{\frac{2mg}{C\rho A}}$$

v_T = terminal velocity
mg = weight

equation 3

	Drag coefficient
Ice cream cone	0.34
Bowl of petunias	0.41
Can of soup	0.88
Dinner plate	1.11
Parachute	1.35

Drag coefficients
Based on approximations of shape



The drag coefficient *C* is 0.49 and the air density *ρ* is 1.1 kg/m^3. What is the skydiver's terminal velocity?

$$v_T = \sqrt{\frac{2mg}{C\rho A}}$$

$$v_T = \sqrt{\frac{2(650\text{ N})}{(0.49)(1.1\text{ kg/m}^3)(0.90\text{ m}^2)}}$$

$$v_T = 52\text{ m/s}$$

5.31 - Interactive summary problem: helicopter pilot

In the scenario to the right, the helicopter has a mass of 1710 kg. The simulation starts paused with the helicopter 25.0 meters above the ground. When you press GO the helicopter is moving with an initial velocity of −3.60 m/s. You want to determine the appropriate lift force so that it achieves zero velocity at a height of 5.00 meters. If you calculate correctly, it will attach itself to the large crate and lift it. What force should the helicopter blades supply to allow you to accomplish this?

To solve this problem, you will need to first determine the appropriate acceleration using a motion equation, and then calculate the net force required. Hint: Make sure you consider both of the forces acting on the helicopter! Enter the force to the nearest 100 N.

When you successfully pick up the crate, notice that the acceleration decreases because the same net force is now applied to an object (helicopter plus crate) with increased mass.

If you have difficulty solving this, consult the section on motion equations and Newton's second law.

5.32 - Gotchas

An object has a speed of 20 km/h. It swerves to the left but maintains the same speed. Was a force involved? Yes. A change in speed **or** direction is acceleration, and acceleration requires a force.

An object is moving. A net force must be acting on it. No. Only if the object is accelerating (changing speed or direction) is there a net force. Constant velocity means there is no net force.

No acceleration means no forces are present. Close. No acceleration means no **net** forces. There can be a balanced set of forces and no acceleration.

"*I weigh 70 kilograms."* False. Kilograms measure mass, not weight.

"*I weigh the same on Jupiter as I do on Mars."* Not unless you dieted (lost mass) as you traveled from Jupiter to Mars. Weight is gravitational force, and Mars exerts less gravitational force.

"*My mass is the same on Jupiter and Mars."* Yes.

The normal force is the response force to gravity. This is too specific of a definition. The normal force appears any time two objects are brought in contact. It is not limited to gravity. For instance, if you lean against a wall, the force of the wall on you is a normal force. If you stand on the ground, the normal force of the ground is a response force to gravity.

"I push against a wall with a force of five newtons. The wall pushes back with the same force." Close, but it is better to say, "The same amount (magnitude) of force but in the opposite direction."

"I pull on the Earth with the same amount of gravitational force that the Earth exerts on me." True. You are an action-reaction pair.

5.33 - Summary

Force, and Newton's laws which describe force, are fundamental concepts in the study of physics. Force can be described as a push or pull. It is a vector quantity that is measured in newtons (1 N = 1 kg·m/s^2). Net force is the vector sum of all the external forces on an object.

Free-body diagrams depict all the external forces on an object. Even though the forces may act on different parts of the object, free-body diagrams are drawn so that the forces are shown as being applied at a single point.

Newton's first law states that an object maintains a constant velocity (including remaining at rest) until a net force acts upon it.

Mass is the property of an object that determines its resistance to a change in velocity, and it is a scalar, measured in kilograms. Mass should not be confused with weight, which is a force caused by gravity, directed toward the center of the Earth.

Newton's second law states that the net force on an object is equal to its mass times its acceleration.

Newton's third law states that the forces that two bodies exert on each other are always equal in magnitude and opposite in direction.

The normal force is a force that occurs when two objects are in direct contact. It is always directed perpendicular to the surface of contact.

Tension is a force exerted by a means of connection such as a rope, and the tension force always pulls on the bodies to which the rope is attached.

Friction is a force that resists the sliding motion of two objects in direct contact. It is proportional to the magnitude of the normal force and varies according to the composition of the objects.

Equations

$\text{weight} = mg$

Newton's second law

$\Sigma\mathbf{F} = m\mathbf{a}$

Newton's third law

$\mathbf{F}_{ab} = -\mathbf{F}_{ba}$

Static friction

$f_{s,max} = \mu_s F_N$

Kinetic friction

$f_k = \mu_k F_N$

Hooke's law

$F_s = -kx$

Static friction is the term for friction when there is no relative motion between two objects. It balances any applied pushing force that tends to slide the body, up to a maximum determined by the normal force and the coefficient of static friction between the two objects, μ_s. If the applied pushing force is greater than the maximum static friction force, then the object will move.

Once an object is in motion, kinetic friction applies. The force of kinetic friction is determined by the magnitude of the normal force multiplied by μ_k, the coefficient of kinetic friction between the two objects.

Hooke's law describes the force that a spring exerts when stretched or compressed away from its equilibrium position. The force increases linearly with the displacement from the equilibrium position. The equation for Hooke's law includes k, the spring constant; a value that depends on the particular spring. The negative sign indicates that the spring force is a restoring force that points in the opposite direction as the displacement, that is, it resists both stretching and compression.

Air resistance, or drag, is a force that opposes motion through a fluid such as air. Drag increases as speed increases. Terminal velocity is reached when the drag force on a falling object equals its weight, so that it ceases to accelerate.

Conceptual Problems

C.1 Can an object move if there is no net force on it? Explain.

○ Yes ○ No

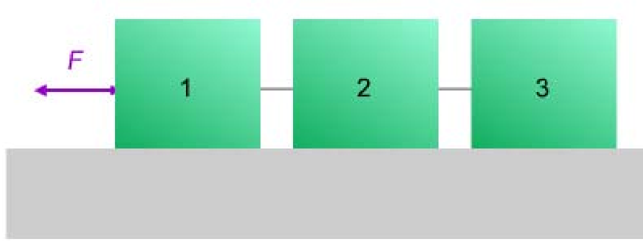
C.2 Suppose you apply a force of 1 N to block A and a force of 2 N to block B. Does it follow that block B has twice the acceleration of block A? Justify your answer using Newton's second law.

○ Yes ○ No

C.3 When a brick rests on a flat, stationary, horizontal table, there is an upward normal force on it from the table. Explain why the brick does not accelerate upward in response to this force.

C.4 A rocket in space can change course with its engines. Since in empty space there is nothing for the exhaust gases to push on, how can it accelerate?

C.5 Blocks 1 and 2, and 2 and 3 are connected by two identical thin wires. All three blocks are resting on a frictionless table. Block 1 is pulled by a constant force and all three blocks accelerate equally in a line, with block 1 leading. Are the tensions in the two wires the same or different? If the tensions are different, which has the larger magnitude? Why?



i. Tensions are the same
ii. Greater between blocks 1 and 2
iii. Greater between blocks 2 and 3

C.6 Two blocks of different mass are connected by a massless rope which goes over a massless, frictionless pulley. The rope is free to move, and both of the blocks hang vertically. What is the magnitude of the tension in the rope?

i. The weight of the heavier block
ii. The weight of the lighter block
iii. Their combined weight
iv. A value between the two weights
v. zero

C.7 Why is the frictional force proportional to the normal force, and not weight?

C.8 A college rower can easily push a small car along a flat road, but she cannot lift the car in the air. Since the mass of the car is constant, how can you explain this discrepancy?

C.9 Without friction, you would not be able to walk along a level sidewalk. Why? (Imagine being stranded in the middle of an ice rink, wearing shoes made of ice.)

C.10 If an acrobat who weighs 800 N is clinging to a vertical pole using only his hands, neither moving up nor down, can we determine the coefficient of static friction between his hands and the pole? Explain your answer.

○ Yes ○ No

C.11 State two reasons why it is easier to push a heavy object down a hill than it is to push that same object across a flat, horizontal surface.

C.12 One end of a spring is attached firmly to a wall, and a block is attached to the other end. When the spring is fully compressed, it exerts a force *F* on the block, and when the spring is fully extended, the force it exerts on the block is −*F*. What is the force of the spring on the block at (a) equilibrium (neither compressed nor stretched), (b) halfway between maximum stretch and equilibrium, and (c) halfway between maximum compression and equilibrium? Carefully consider the signs in your answer, which indicate direction, and express your answers in terms of *F*.

(a)
i. F
ii. -F
iii. 0
iv. -F/2
v. F/2

(b)
i. F
ii. -F
iii. 0
iv. -F/2
v. F/2

(c)
i. F
ii. -F
iii. 0
iv. -F/2
v. F/2

C.13 From the example of the falling cat, we see that the cross-sectional area of a falling object affects its terminal velocity. Does an object's mass also affect its terminal velocity? Why or why not?

○ Yes ○ No

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following question. If the net force on the helicopter is zero, what must the helicopter be doing?

i. Rising
ii. Falling
iii. Staying still
iv. It can be doing any of these

Section 2 - Newton's first law

2.1 An airplane of mass 2867 kg flies at a constant horizontal velocity. The force of air resistance on it is 2225 N. What is the net force on the plane (magnitude and direction)?

i. 0 N (direction does not matter)
ii. 2225 N opposing the motion
iii. 642 N at 30° below horizontal
iv. It is impossible to say

2.2 Three children are pulling on a toy that has a mass of 1.25 kg. Child A pulls with force (5.00, 6.00, 2.50) N. Child B pulls with force (0.00, 9.20, 2.40) N. With what force must Child C pull for the toy to remain stationary?

(__________ , __________ , __________) N

Section 4 - Gravitational force: weight

4.1 A weightlifter can exert an upward force of 3750 N. If a dumbbell has a mass of 225 kg, what is the maximum number of dumbbells this weightlifter could hold simultaneously if he were on the Moon? (The Moon's acceleration due to gravity is approximately 0.166 times freefall acceleration on Earth.) The weightlifter cannot pick up a fraction of a dumbbell, so make sure your answer is an integer.

__________ dumbells

4.2 (a) How much does a 70.0 kg person weigh on the Earth? (b) How much would she weigh on the Moon (g_{moon} = 0.166g)? (c) How much would she weigh on a neutron star where g_{star} = 1.43×10^{11}g?

(a) __________ N
(b) __________ N
(c) __________ N

4.3 A dog weighs 47.0 pounds on Earth. (a) What is its weight in newtons? (One newton equals 0.225 pounds.) (b) What is its mass in kilograms?

(a) __________ N
(b) __________ kg

4.4 A dog on Earth weighs 136 N. The same dog weighs 154 N on Neptune. What is the acceleration due to gravity on Neptune?

__________ m/s^2

Section 5 - Newton's second law

5.1 When empty, a particular helicopter of mass 3770 kg can accelerate straight upward at a maximum acceleration of 1.34 m/s^2. A careless crewman overloads the helicopter so that it is just unable to lift off. What is the mass of the cargo?

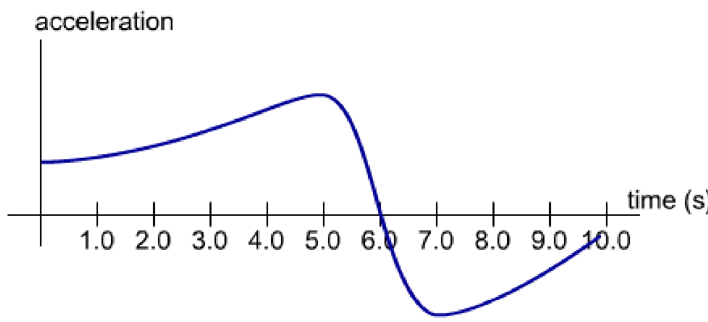
__________ kg

5.2 A 0.125 kg frozen hamburger patty has two forces acting on it that determine its horizontal motion. A 2.30 N force pushes it to the left, and a 0.800 N force pushes it to the right. (a) Taking right to be positive, what is the net force acting on it? (b) What is its acceleration?

(a) __________ N
(b) __________ m/s^2

5.3 In the illustration, you see the graph of an object's acceleration over time. (a) At what moment is it experiencing the most positive force? (b) The most negative force? (c) Zero force?

(a) __________ s
(b) __________ s
(c) __________ s



5.4 The net force on a boat causes it to accelerate at 1.55 m/s^2. The mass of the boat is 215 kg. The same net force causes another boat to accelerate at 0.125 m/s^2. (a) What is the mass of the second boat? (b) One of the boats is now loaded on the other, and the same net force is applied to this combined mass. What acceleration does it cause?

(a) __________ kg
(b) __________ m/s^2

5.5 A flea has a mass of 4.9×10^{-7} kg. When a flea jumps, its rear legs act like catapults, accelerating it at 2400 m/s^2. What force do the flea's legs have to exert on the ground for a flea to accelerate at this rate?

__________ N

5.6 An extreme amusement park ride accelerates its riders upward from rest to 50.0 m/s in 7.00 seconds. Ignoring air resistance, what average upward force does the seat exert on a rider who weighs 1120 N?

__________ N

5.7 A giant excavator (used in road construction) can apply a maximum vertical force of 2.25×10^5 N. If it can vertically accelerate a load of dirt at 0.200 m/s^2, what is the mass of that load? Ignore the mass of the excavator itself.

__________ kg

5.8 In moving a standard computer mouse, a user applies a horizontal force of 6.00×10^{-2} N. The mouse has a mass of 125 g. (a) What is the acceleration of the mouse? Ignore forces like friction that oppose its motion. (b) Assuming it starts from rest, what is its speed after moving 0.159 m across a mouse pad?

(a) __________ m/s^2
(b) __________ m/s

5.9 A solar sail is used to propel a spacecraft. It uses the pressure (force per unit area) of sunlight instead of wind. Assume the sail and its spacecraft have a mass of 245 kg. If the sail has an area of 62,500 m^2 and achieves a velocity of 8.93 m/s in 12.0 hours starting from rest, what pressure does light of the Sun exert on the sail? To simplify the problem, ignore other forces acting on the spacecraft and assume the pressure is constant even as its distance from the Sun increases.

__________ N/m^2

5.10 A tennis player strikes a tennis ball of mass 56.7 g when it is at the top of the toss, accelerating it to 68.0 m/s in a distance of 0.0250 m. What is the average force the player exerts on the ball? Ignore any other forces acting on the ball.

__________ N

5.11 There are two forces acting on a box of golf balls, $\mathbf{F}_1$ and $\mathbf{F}_2$. The mass of the box is 0.750 kg. When the forces act in the same direction, they cause an acceleration of 0.450 m/s^2. When they oppose one another, the box accelerates at 0.240 m/s^2 in the direction of $\mathbf{F}_2$. (a) What is the magnitude of $\mathbf{F}_1$? (b) What is the magnitude of $\mathbf{F}_2$?

(a) _________ N
(b) _________ N

5.12 A chain of roller coaster cars moving horizontally comes to an abrupt stop and the passengers are accelerated by their safety harnesses. In one particular car in the chain, the car has a mass M = 122 kg, the first passenger has a mass m_1 = 55.2 kg, and the second passenger has a mass m_2 = 68.8 kg. If the chain of cars slows from 26.5 m/s to a stop in 4.73 s, calculate the average magnitude of force exerted by their safety harnesses on (a) the first passenger and (b) the second passenger.

(a) _________ N
(b) _________ N

5.13 Cedar Point's Top Thrill Dragster Strata-Coaster in Ohio, the fastest amusement park ride in the world as of 2004, can accelerate its riders from rest to 193 km/h in 4.00 seconds. (a) What is the magnitude of the average acceleration of a rider? (b) What is the average net force on a 45.0 kg rider during these 4.00 seconds? (c) Do people really pay money for this?

(a) _________ m/s^2
(b) _________ N
(c) ○ Yes ○ No

5.14 The leader is the weakest part of a fly-fishing line. A given leader can withstand 19 N of force. A trout when caught will accelerate, taking advantage of slack in the line, and some trout are strong enough to snap the line. Assume that with the line taut and the rod unable to flex further, a 1.3 kg trout is just able to snap this leader. How much time would it take this trout to accelerate from rest to 5.0 m/s if it were free of the line? Note: Trout can reach speeds like this in this interval of time.

_________ s

5.15 A 7.6 kg chair is pushed across a frictionless floor with a force of 42 N that is applied at an angle of 22° downward from the horizontal. What is the magnitude of the acceleration of the chair?

_________ m/s^2

Section 9 - Interactive problem: flying in formation

9.1 Using the simulation in the interactive problem in this section, (a) what is the force required for the red ships to accelerate at the desired magnitude? (b) What force is required for the blue ships?

(a) _________ N
(b) _________ N

Section 10 - Newton's third law

10.1 A 75.0 kg man sits on a massless cart that is on a horizontal surface. The cart is initially stationary and it can move without friction or air resistance. The man throws a 5.00 kg stone in the positive direction, applying a force to it so that it has acceleration +3.50 m/s^2 as measured by a nearby observer on the ground. What is the man's acceleration during the throw, as seen by the same observer? Be careful to use correct signs.

_________ m/s^2

10.2 Two motionless ice skaters face each other and put their palms together. One skater pushes the other away using a constant force for 0.80 s. The second skater, who is pushed, has a mass of 110 kg and moves off with a velocity of −1.2 m/s relative to the rink. If the first skater has a mass of 45 kg, what is her velocity relative to the rink after the push? (Consider any forces other than the push acting on the skaters as negligible.)

_________ m/s

10.3 Two cubic blocks are in contact, resting on a frictionless horizontal surface. The block on the left has a mass of m_L = 6.70 kg, and the block on the right has a mass of m_R = 18.4 kg. A force of magnitude 112 N is applied to the left face of the left block, toward the right but at an upward angle of 30.0° with respect to the horizontal. It causes the left block to push on the right block. What are (a) the direction and (b) the magnitude of the force that the **right** block applies to the **left** block?

(a)
i. To the left at a 30.0° angle down
ii. To the right at a 30.0° angle up
iii. Directly left
iv. Directly right

(b) _________ N

Section 11 - Normal force

11.1 A cup and saucer rest on a table top. The cup has mass 0.176 kg and the saucer 0.165 kg. Calculate the magnitude of the normal force (a) the saucer exerts on the cup and (b) the table exerts on the saucer.

(a) __________ N
(b) __________ N

11.2 Three blocks are arranged in a stack on a frictionless horizontal surface. The bottom block has a mass of 37.0 kg. A block of mass 18.0 kg sits on top of it and a 8 kg block sits on top of the middle block. A downward vertical force of 170 N is applied to the top block. What is the magnitude of the normal force exerted by the bottom block on the middle block?

__________ N

11.3 A 22.0 kg child slides down a slide that makes a 37.0° angle with the horizontal. (a) What is the magnitude of the normal force that the slide exerts on the child? (b) At what angle from the horizontal is this force directed? State your answer as a number between 0 and 90°.

(a) __________ N
(b) __________ °

11.4 A 6.00 kg box is resting on a table. You push down on the box with a force of 8.00 N. What is the magnitude of the normal force of the table on the block?

__________ N

Section 12 - Tension

12.1 An ice rescue team pulls a stranded hiker off a frozen lake by throwing him a rope and pulling him horizontally across the essentially frictionless ice with a constant force. The hiker weighs 1040 N, and accelerates across the ice at 1.10 m/s^2. What is the magnitude of the tension in the rope? (Ignore the mass of the rope.)

__________ N

12.2 During recess, Maria, who has mass 27.0 kg, hangs motionless on the monkey bars, with both hands gripping a horizontal bar. Assume her arms are vertical and evenly support her body. What is the tension in each of her arms?

__________ N

Section 14 - Free-body diagrams

14.1 A tugboat is towing an oil tanker on a straight section of a river. The current in the river applies a force on the tanker that is one-half the magnitude of the force that the tugboat applies. Draw a free-body diagram for the tanker when the force provided by the tugboat is directed (a) straight upstream (b) directly cross-stream, and (c) at a 30° angle to upstream. In your drawing, let the current flow in the "left" direction, and have the tugboat pull "up" when applying a cross-stream force. Label the force vectors.

14.2 A person lifts a 3.60 kg textbook (remember when they were made of paper and so heavy?) with a 52.0 N force at a 60.0° angle from the horizontal. (a) Draw a free-body diagram of the forces acting on the book, ignoring air resistance. Label the forces. (b) Draw a free-body diagram showing the net force acting on the book.

14.3 A 17.0 N force *F* acting on a 4.00 kg block is directed at 30.0° from the horizontal, parallel to the surface of a frictionless ramp. Draw a free-body diagram of the forces acting on the block, including the normal force, and label the forces. Make sure the vectors are roughly proportional to the forces!

14.4 An apple is resting on a table. Draw free-body diagrams for the apple, table and the Earth.

Section 15 - Interactive problem: free-body diagram

15.1 Using simulation in the interactive problem in this section, what are the direction and magnitude of (a) the weight, (b) the normal force, (c) the tension, and (d) the frictional force that meet the stated requirements and give the desired acceleration?

(a) __________ N,
i. Up
ii. Down
iii. Left
iv. Right

(b) __________ N,
i. Up
ii. Down
iii. Left
iv. Right

(c) __________ N,
i. Up
ii. Down
iii. Left
iv. Right

(d) __________ N,
i. Up
ii. Down
iii. Left
iv. Right

Section 17 - Interactive problem: lifting crates

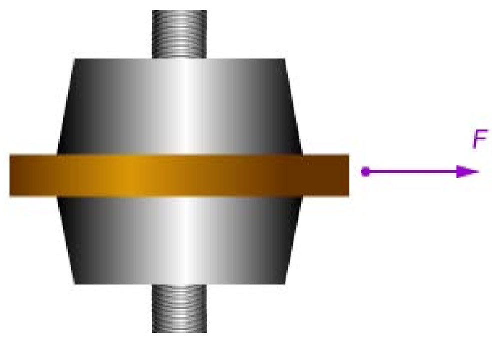
17.1 Use the helicopter in the simulation in the interactive problem in this section to find the crate with the given mass.

i. The leftmost crate
ii. The middle crate
iii. The rightmost crate

Section 19 - Static friction

19.1 A piece of steel is held firmly in the jaws of a vise. A force larger than 3350 N will cause the piece of steel to start to move out of the vise. If the coefficient of friction between the steel and each of the jaws of the vise is 0.825 and each jaw applies an equal force, what is the magnitude of the normal force exerted on the steel by each jaw?

__________ N



19.2 A wooden block of mass 29.0 kg sits on a horizontal table. A wire of negligible mass is attached to the right side of the block and goes over a pulley (also of negligible mass, and frictionless), where it is allowed to dangle vertically. When a mass of 15.5 kg is attached to the dangling wire, the block on the table just barely starts to slide. What is the coefficient of static friction between the block and the table?

____

19.3 The Occupational Safety and Health Administration (OSHA) suggests a minimum coefficient of static friction of μ_s = 0.50 for floors. If Ethan, who has mass of 53 kg, stands passively, how much horizontal force can be applied on him before he will slip on a floor with OSHA's minimum coefficient of static friction?

__________ N

19.4 A block of mass *m* sits on top of a larger block of mass 2*m*, which in turn sits on a flat, frictionless table. The coefficient of static friction between the two blocks is μ_s.What is the largest possible horizontal acceleration you can give the bottom block without the top block slipping?

- ○ $\mu_s g/2$
- ○ $\mu_s g$
- ○ $2\mu_s g$

Section 20 - Kinetic friction

20.1 A 1.0 kg brick is pushed against a vertical wall by a horizontal force of 24 N. If μ_s = 0.80 and μ_k = 0.70 what is the acceleration of the brick?

__________ m/s2

20.2 A firefighter whose weight is 812 N is sliding down a vertical pole, her speed increasing at the rate of 1.45 m/s^2. Gravity and friction are the two significant forces acting on her. What is the magnitude of the frictional force?

__________ N

20.3 A plastic box of mass 1.1 kg slides along a horizontal table. Its initial speed is 3.9 m/s, and the force of kinetic friction opposes its motion, causing it to stop after 3.1 s. What is the coefficient of kinetic friction between the block and the table?

____

20.4 An old car is traveling down a long, straight, dry road at 25.0 m/s when the driver slams on the brakes, locking the wheels. The car comes to a complete stop after sliding 215 m in a straight line. If the car has a mass of 755 kg, what is the coefficient of kinetic friction between the tires and the road?

____

20.5 A rescue worker pulls an injured skier lying on a toboggan (with a combined mass of 127 kg) across flat snow at a constant speed. A 2.43 m rope is attached to the toboggan at ground level, and the rescuer holds the rope taut at shoulder level. If the rescuer's shoulders are 1.65 m above the ground, and the tension in the rope is 148 N, what is the coefficient of kinetic friction between the toboggan and the snow?

____

20.6 You are trying to move a chair of mass 29.0 kg by pushing it horizontally along the floor, but it is not sliding very easily. The coefficient of kinetic friction between the chair and the floor is 0.700. (a) If you push the chair with a horizontal force, what is the magnitude of the minimum force that will keep the chair moving at a constant velocity? (b) If the force on the chair is directed 30.0 degrees up from the horizontal, what is the magnitude of the minimum force that will keep the chair moving at a constant velocity? (c) What if the angle is 75.0 degrees from the horizontal?

(a) __________ N
(b) __________ N
(c) __________ N

Section 24 - Interactive problem: forces on a sliding block

24.1 Using the simulation in the interactive problem in this section, what are the direction and magnitude of (a) the weight, (b) the normal force, (c) the frictional force, and (d) the tension that meet the stated requirements and give the desired acceleration?

(a) __________ N,
i. Straight up
ii. Straight down
iii. Up the plane
iv. To the right

(b) __________ N,
i. Straight up
ii. Straight down
iii. Up and to the left
iv. To the right

(c) __________ N,
i. Down the plane
ii. Up the plane
iii. To the right
iv. To the left

(d) __________ N,
i. Down the plane
ii. Up the plane
iii. Perpendicular to the plane
iv. Straight down

Section 25 - Sample problem: moving down a frictionless plane

25.1 A child sits on a freshly oiled, straight stair rail that is effectively frictionless and slides down it. She has a mass of 25 kg, and the rail makes an angle of 40° above the ground. If she slides 4.0 m before reaching the bottom, what is her speed there?

__________ m/s

25.2 A shipping container is hauled up a roller ramp that is effectively frictionless at a constant speed of 2.10 m/s by a 2250 N force that is parallel to the ramp. If the ramp is at a 24.6° incline, what is the container's mass?

__________ kg

Section 28 - Hooke's law and spring force

28.1 A 10.0 kg mass is placed on a frictionless, horizontal surface. The mass is connected to the end of a horizontal compressed spring which has a spring constant 339 N/m. When the spring is released, the mass has an initial, positive acceleration of 10.2 m/s^2. What was the displacement of the spring, as measured from equilibrium, before the block was released? Watch the sign of your answer.

__________ m

28.2 Consider a large spring, hanging vertically, with spring constant k = 3220 N/m. If the spring is stretched 25.0 cm from equilibrium and a block is attached to the end, the block stays still, neither accelerating upward nor downward. What is the mass of the block?

__________ kg

28.3 A spring with spring constant k = 15.0 N/m hangs vertically from the ceiling. A 1.20 kg mass is attached to the bottom end of the spring, and allowed to hang freely until it becomes stationary. Then, the mass is pulled downward 10.0 cm from its resting position and released. At the moment of its release, what is (a) the magnitude of the mass's acceleration and (b) the direction? Ignore the mass of the spring.

(a) __________ m/s^2
(b) i. Downward
ii. Upward

28.4 A 5.00 kg wood cube rests on a frictionless horizontal table. It has two springs attached to it on opposite faces. The spring on the left has a spring constant of 55.0 N/m, and the spring on the right has a spring constant of 111 N/m. Both springs are initially in their equilibrium positions (neither compressed nor stretched). The block is moved toward the left 10.0 cm, compressing the left spring and stretching the right spring. (a) Calculate the resulting net force on the block. (b) Calculate the the initial acceleration of the block when it is released. Use the convention that to the right is positive and to the left is negative.

(a) __________ N
(b) __________ m/s^2

28.5 A man steps on his bathroom scale and obtains a reading of 243 lb. The spring in the scale is compressed by a displacement of −0.0590 inches. Calculate the value of its spring constant in (a) pounds per inch (b) newtons per meter.

(a) __________ lbs/in
(b) __________ N/m

Section 30 - Air resistance

30.1 A parachutist and her parachuting equipment have a combined mass of 101 kg. Her terminal velocity is 5.30 m/s with the parachute open. Her parachute has a cross-sectional area of 35.8 m^2. The density of air at that altitude is 1.23 kg/m^3. What is the drag cofficient of the parachutist with her parachute?

____

30.2 A tennis ball of mass 57.0 g is dropped from the observation deck of the Empire State building (369 m). The tennis ball has a cross-sectional area of 3.50×10^{-3} m^2 and a drag coefficient of 0.600. Using 1.23 kg/m^3 for the density of air, (a) what is the speed of the ball when it hits the ground? (b) What would be the final speed of the ball if you did not include air resistance in your calculations? Remember to use the appropriate sign when answering the question.

(a) __________ m/s
(b) __________ m/s

Section 31 - Interactive summary problem: helicopter pilot

31.1 Use the simulation in the interactive problem in this section to calculate the lift force required for the helicopter to pick up the crate.

__________ N

Additional Problems

A.1 A woman is pushing a box along the ground. She is exerting a horizontal force of 156 N on the box. The box has a mass of 84.0 kg, and has an opposing horizontal frictional force of 135 N. What is the coefficient of kinetic friction between the box and the ground?

____

A.2 A rocket skateboard of mass 3.00 kg has an initial velocity of (5.00, 2.50) m/s. A net force of (3.40, 2.67) N is applied to it for 10.0 s. At the end of this time, what is its velocity?

(__________ , __________) m/s

A.3 A popular ride at an amusement park involves standing in a cylindrical room. The room rotates and the wall presses against the rider as the floor drops down until the rider is no longer touching it. Annabel has a mass of 46.0 kg and the coefficient of static friction between the wall and her is 0.450. What is the minimum normal force exerted on her from the wall that will keep Annabel from sliding down?

__________ N

A.4 A large crate has a mass of 214 kg. A horizontal force of positive 196 N is applied to it, causing it to accelerate at 0.130 m/s^2 horizontally. What is the force of friction opposing the motion of the crate? Use the correct sign to indicate the direction of the force of friction. It is the sole force opposing the crate's motion in this direction.

__________ N

A.5 A ball moving through a special sticky fluid encounters a drag force whose magnitude, in newtons, is proportional to the fourth power of its velocity, expressed in meters per second. That is, $F_D = Av^4$. (a) What are the units of the coefficient "A"? Express your answer using the SI base units of kilograms, meters, and seconds. (b) If a ball of mass m is dropped from rest into a deep container of the sticky fluid, find an expression for the terminal velocity that it reaches.

(a) ○ kg·m/s ○ kg·s^2/m^3 ○ kg·m^2/s^3 ○ kg·s/m^2

(b) Submit answer on paper.

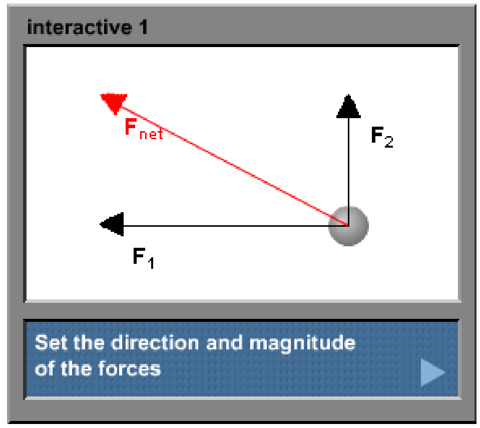
A.6 A 2.0 kg mass rests on a frictionless wedge that has an acceleration of 15 m/s^2 to the right. The mass remains stationary relative to the wedge, moving neither higher nor lower. (a) What is the angle of inclination, θ, of the wedge? (b) What is the magnitude of the normal force exerted on the mass by the incline? (c) What would happen if the wedge were given a greater acceleration?

(a) __________ °
(b) __________ N
(c) i. The block would slide up the wedge
ii. The block would remain at the same location on the wedge
iii. The block would slide down the wedge

A.7 Consider a road bicycle being pedaled along a level street. The foot pedals drive a chain that supplies the force of the biker's legs to the rear wheel, rotating it. The bike accelerates forward, and both tires roll without slipping along the road. (a) In which direction, forward or backward, does friction exert a force on the rear tire? (b) In which direction does friction act on the front tire? (c) Explain your answers.

(a) i. Forward
ii. Backward
(b) i. Forward
ii. Backward
(c) Submit answer on paper

6.0 - Introduction



Now, you will get some additional practice applying Newton's laws. More specifically, you will use them in situations where multiple forces are acting on a single object.

If the application of multiple forces results in a net force acting on an object, it accelerates. On the other hand, if the forces acting on it sum to zero in every dimension, the result is equilibrium. The object does not accelerate; it either maintains a constant velocity, or remains stationary. (Forces can also cause an object to rotate, but rotational motion is a later topic in mechanics.)

Equilibrium is an important topic in engineering. The school buildings you study in, the bridges you travel across − all such structures require careful design to ensure that they remain in equilibrium.

The simulation on the right will help you develop an understanding for how forces in different directions combine when applied to an object. The 5.0 kg ball has two forces acting on it, $\mathbf{F}_1$ and $\mathbf{F}_2$. They act on it as long as the ball is on the screen.

You control the direction and magnitude of each force. In the simulation, you set a force vector's direction and magnitude by dragging its arrowhead; You will notice the angles are restricted to multiples of $90°$. You can also adjust the magnitude of each vector with a controller in the control panel. The net force is shown in the simulation; it is the vector sum of $\mathbf{F}_1$ and $\mathbf{F}_2$.

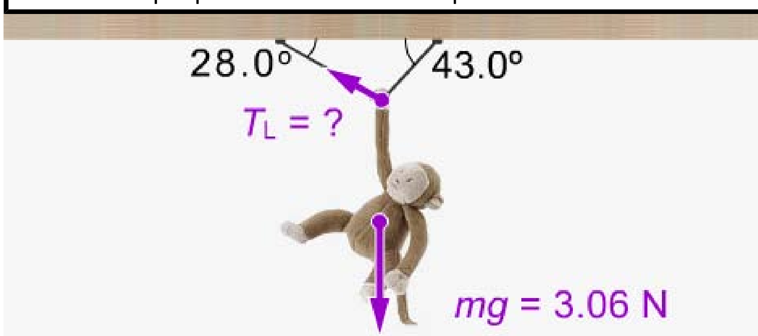
You can check the box "Display vectors head to tail" if you would like to see them graphically combined in that fashion. Press GO to start the simulation and set the ball moving in response to the forces on it.

Here are some challenges for you. First, set the forces so that the ball does not move at all when you press GO. The individual forces must be at least 10 newtons, so setting them both to zero is not an option!

Next, hit each of the three animated targets. The center of one is directly to the right of the ball and the center of another is at a $45°$ angle above the horizontal from the ball. Set the individual vectors and press GO to hit the center of each target in turn.

The target to the left is at a $150°$ angle. It is the "extra credit" target. Determining the correct ratio of vectors will require a little thought. We allow for rounding with this target; if you set one of the vectors to 10 N, you can solve the problem by setting the other one to the appropriate closest integer value.

6.1 - Sample problem: a mass on ropes



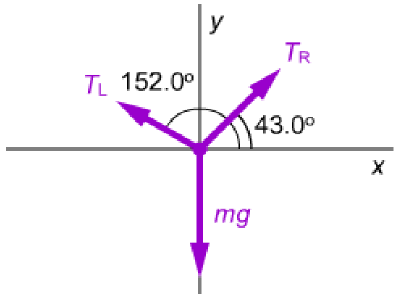
The monkey hangs without moving in the configuration you see here. What is the magnitude of the tension in the left rope?

Since it is stationary, the monkey is not accelerating, which means no net force is acting on it.

This section shows you a useful technique for solving problems that involve multiple forces acting on a single object. To calculate the overall force on an object like the rope, the x and y components of each force need to be determined. Since there is no acceleration in any direction, there is no net force along any dimension. Two equations can be developed: The sum of the x components equals zero, and the sum of the y components equals zero.

We will use a consistent approach to solving multi-force problems. First we draw a free-body diagram to help us identify the variables and the force components. Then we state the variables and their values when they are known. Next, we use Newton's second law, relating the net force to the acceleration and the mass of objects in the problem. Finally, we perform the algebraic and mathematical steps needed to solve the problem.

Draw a free-body diagram



Variables

Since some of the forces do not lie along either the x or y axis, we use trigonometry to calculate the components of each force. For consistency, we do this even for the weight, which points straight down in the negative y direction.

	x component	y component
tension to left	$T_L \cos 152.0°$	$T_L \sin 152.0°$
tension to right	$T_R \cos 43.0°$	$T_R \sin 43.0°$
weight	0 N	$mg \sin 270°$
weight	$mg = 3.06$ N	

What is the strategy?

1. Draw a free-body diagram of the forces on the monkey.
2. Set the net force equal to zero in each dimension.
3. This gives two equations with two unknowns. Use algebra to solve for the left tension force.

Physics principles and equations

Newton's second law

$$\Sigma \mathbf{F} = m\mathbf{a}$$

When the acceleration is zero, the forces along every dimension must sum to zero.

Step-by-step solution

We first apply the equilibrium condition in the x dimension to relate the tension forces in the two ropes.

Step	Reason
1. $\Sigma \mathbf{F}_x = 0$	no acceleration in x dimension
2. $T_L \cos 152.0° + T_R \cos 43.0° = 0$	add x components
3. $T_L(-0.883) + T_R(0.731) = 0$	trigonometry
4. $T_R = T_L(1.21)$	solve for T_R

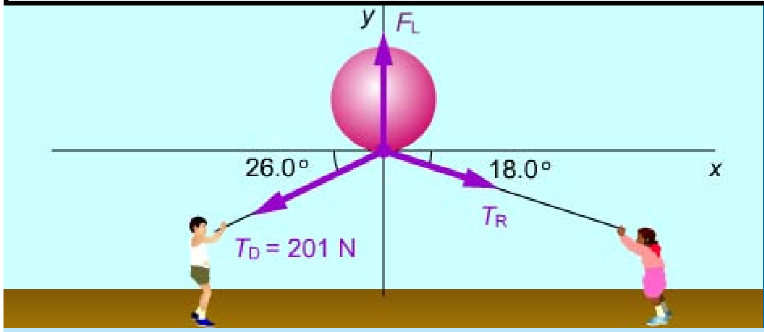
We have one relationship between the two tensions. Now we use equilibrium applied in the y dimension to develop a second equation. The y components of the rope tensions balance the monkey's weight.

Step	Reason
5. $\Sigma \mathbf{F}_y = 0$	no acceleration in the y dimension
6. $T_L \sin 152.0° + T_R \sin 43.0° + (3.06\text{ N}) \sin 270° = 0$	add y components
7. $T_L(0.469) + T_R(0.682) = 3.06$ N	trigonometry

We now have two equations in two unknowns. We substitute one equation into the other and solve for the tension in the left rope.

Step	Reason
8. $T_L(0.469) + T_L(1.21)(0.682) = 3.06$ N	substitute equation 4 into equation 7
9. $T_L(1.29) = 3.06$ N	simplify
10. $T_L = 2.37$ N	division

6.2 - Interactive checkpoint: helium balloon



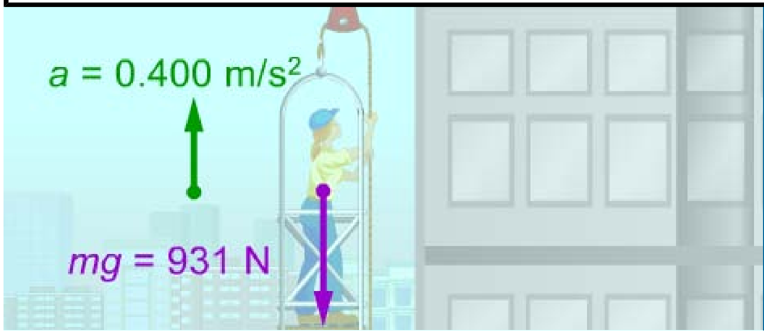
A pair of strings holds a helium balloon in place. Dan and Rosa are each holding onto one of the strings. There is no wind, but the atmosphere provides a vertical lift force F_L on the balloon. The balloon's weight is negligible. What is the magnitude of the tension in Rosa's string? What is the magnitude of F_L?

Answer:

T_R = ______ N

F_L = ______ N

6.3 - Sample problem: pulling up a scaffold



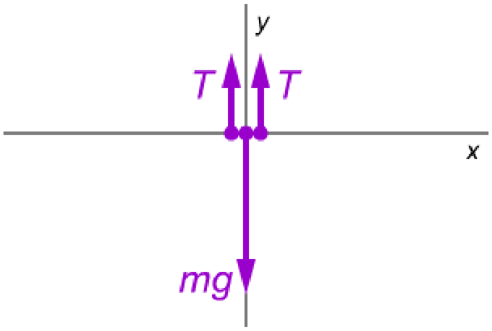
What is the magnitude of the tension in the rope above the window washer's hands?

If you have ever worked in a skyscraper, you have probably seen window washers raising and lowering themselves and their scaffolding. Let's say a particularly energetic window washer is accelerating herself upward, as shown above. You are asked to find the amount of tension in one of the two ropes.

Ropes and pulleys are fairly common in mechanics problems. We will always assume that tension is transmitted with a change in direction but no change in magnitude in a rope that goes around a pulley. You may often be told to draw these conclusions when a problem states that the pulley is assumed to be frictionless and massless, or the rope is massless and does not stretch.

The drawing above shows the weight of the combination of the scaffolding and the window washer, which is 931 N. To apply Newton's second law, we need to compute the mass of the system (window washer plus scaffolding) from its weight.

Draw a free-body diagram



Variables

weight	$-mg = -931$ N
tension in left rope	T
tension in right rope	T
acceleration	$a = 0.400$ m/s^2
mass	$m = 931$ N / 9.80 m/s^2 = 95.0 kg

All the forces are in the vertical direction, which simplifies our work. The weight is directed down, which means we will treat it as a negative quantity in our equations. We treat the upward tensions as positive quantities.

What is the strategy?

1. Draw a free-body diagram of the forces on the scaffolding.
2. Calculate the net force on the scaffolding. In this case, the forces are all vertical, so we only need to consider forces along the y axis.
3. Use Newton's second law to find the tension force.

Physics principles and equations

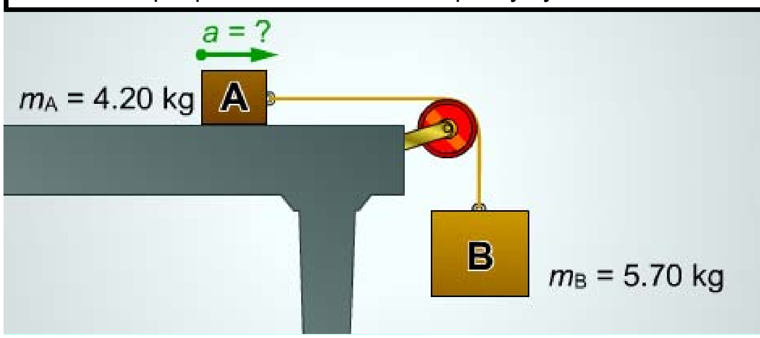
Newton's second law

$$\Sigma\mathbf{F} = m\mathbf{a}$$

Step-by-step solution

Step	Reason
1. $\Sigma\mathbf{F} = m\mathbf{a}$	Newton's second law
2. $T + T + (-mg) = ma$	net force along y axis, down is negative
3. $2T = mg + ma$	rearrange
4. $2T = 931\text{ N} + (95.0\text{ kg})(0.400\text{ m/s}^2)$	enter values
5. $T = 485\text{ N}$	solve for T

6.4 - Sample problem: blocks and a pulley system



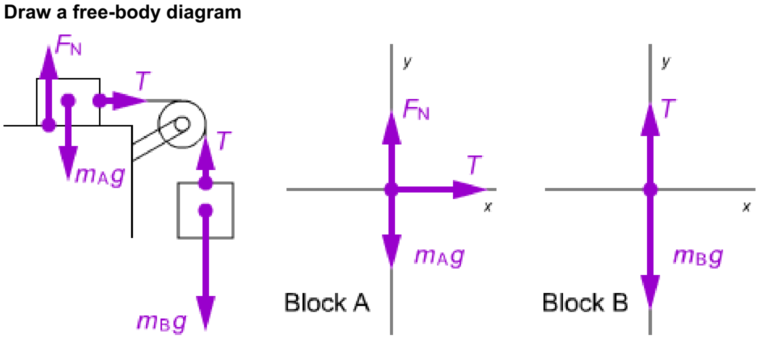
What is the acceleration of the block on the frictionless table?

Two blocks, labeled A and B, are connected by a rope running over a pulley, as shown above. The rope and pulley are massless, the rope does not stretch, and the pulley and the table are friction-free. You are asked to find the acceleration of block A, on the table. This requires **two** free-body diagrams, one for each block.

The magnitudes of the tension forces exerted on each block by the rope are the same; the amount of tension does not vary within a rope. Because the rope connects the blocks, they accelerate at the same rate (but in different directions).

In problems like this, you have to inspect the situation to assign the correct sign to the acceleration of each block. Since the acceleration of the block on the table is to the right, we use the usual convention that it is positive. Since the acceleration of the falling block is downward, we treat it as negative.

Draw a free-body diagram



Variables

Block A on table

	x component	y component
tension	T	0 N
weight	0 N	$-m_A g$
normal force	0 N	F_N
acceleration	a	0 m/s
mass	$m_A = 4.20$ kg	

Falling block B

tension	0 N	T
weight	0 N	$-m_B g$
acceleration	0 m/s	$-a$
mass	$m_B = 5.70$ kg	

What is the strategy?

1. Draw a free-body diagram for each block.
2. Calculate the net force on each block. Block A moves only in the horizontal direction, so we can ignore the vertical forces on it. Block B moves only vertically, and there are no horizontal forces on it.
3. Use Newton's second law for each block to find two expressions for the tension force in the rope, and set the expressions equal to find the acceleration.

Physics principles and equations

Newton's second law relates the net force and acceleration. Block A moves only horizontally, so we will consider only the x direction for it; similarly, we consider only the y direction for block B.

$$\Sigma \mathbf{F} = m\mathbf{a}$$

Step-by-step solution

We start by considering block A, and use Newton's second law to find an equation that gives an expression for the tension force in the rope.

Step	Reason
1. $\Sigma \mathbf{F}_x = m_A \mathbf{a}_x$	Newton's second law applied to A
2. $T = m_A a$	tension is net force on A

Then we apply Newton's second law to block B to find another expression for the tension. Since block B falls, we assign its acceleration a negative value.

Step	Reason
3. $\Sigma \mathbf{F}_y = m_B \mathbf{a}_y$	Newton's second law applied to B
4. $T + (-m_B g) = m_B(-a)$	net force is sum of tension and weight
5. $T = m_B g - m_B a$	solve for T

We set the two expressions for the tension equal. The rest is algebra.

Step	Reason
6. $m_A a = m_B g - m_B a$	set tensions equal, from 2 and 5
7. $a = (m_B g)/(m_A + m_B)$	solve for a
8. $a = \dfrac{(5.70\text{ kg})(9.80\text{ m/s}^2)}{4.20\text{ kg} + 5.70\text{ kg}}$	enter values
9. $a = 5.64$ m/s^2	evaluate

The steps above determine the magnitude of the acceleration. Since the question asked for the acceleration of the block on the table, the full answer is 5.64 m/s^2 to the right.

6.5 - Interactive problem: mountain rescue

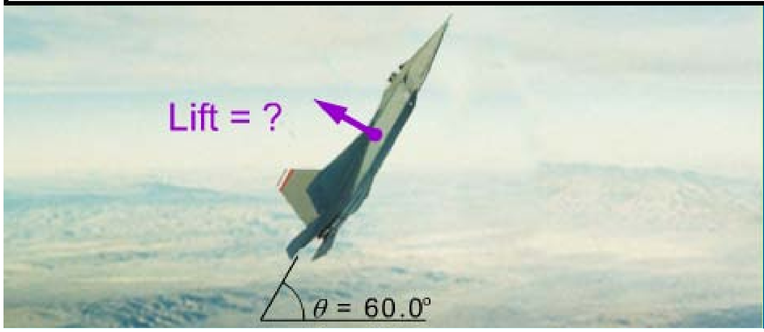
You are in peril! You are hiking solo, with a basket of food and supplies. You have just reached a ledge and the basket is at the bottom, attached to a rope you use to haul up supplies. A hungry grizzly bear has appeared and you need to pull the basket up as quickly as possible.

The basket is attached to a rope that passes through a carabiner (a metal ring that acts as a pulley through which the rope can run freely). You have to pull this rope with a constant force to rescue your food. The rope leaves the carabiner at a $16.0°$ angle from the horizontal. The carabiner is fastened to the rock wall with a chock, a device that is wedged into a crack in the wall. The carabiner is frictionless and the massless rope does not stretch.

The basket has a mass of 18.0 kg. A **horizontal** force on the chock that exceeds 175 N will pull it loose. (Climbers are trained to consider the effects of forces pulling in different directions on gear like chocks.) You need to calculate the maximum force you can apply to the rope to lift the basket up. If the force is too great, the chock will pull loose. If it is too small, the bear will have time to reach the food. Calculate the pulling force on the rope that will result in a horizontal force that just pulls the chock loose. Then reduce the pulling force by 2 N, and you will succeed. Enter the force value you calculate to the nearest newton and press GO to start the simulation. Press RESET to try again.

If you have difficulty solving this, draw a free-body diagram of the forces on the chock, and review the problems in previous sections that deal with forces applied at angles.

6.6 - Sample problem: airplane at constant velocity

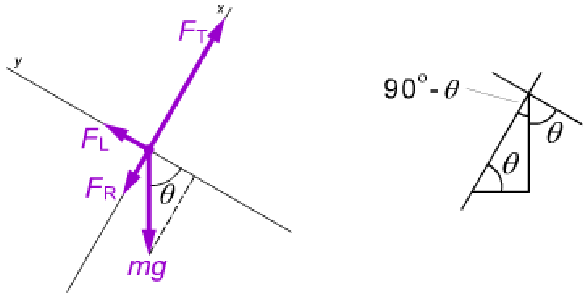


The plane flies at an angle of 60° and is not accelerating. Its weight is 2.60×10^5 N and its engine thrust is 3.00×10^5 N. What is the amount of lift force on its wings?

Above, you see a jet airplane flying through the air at an angle. It travels at a constant velocity; this means it is not accelerating and no net force is acting on the plane along any dimension.

The forces acting on the plane are its weight, the force provided by its engine, called *thrust*, the *drag* force from air resistance, and the *lift* force from the wings. The lift force acts perpendicular to the surface of the wings. How much lift force do the plane's wings provide to keep it aloft?

Draw a free-body diagram



There are four forces on the plane: its weight, the engine thrust, air resistance, and the lift force of the wings. Orienting the axes so that the x axis lies along the length of the airplane will align three of these forces (thrust, air resistance, and lift) along an axis.

With this orientation of the axes, the weight vector makes an angle θ with the y axis. The right-hand diagram above shows that the angle θ in the free-body diagram is the same as the angle at which the plane flies.

Variables

	x component	y component
weight	$-mg \sin \theta$	$-mg \cos \theta$
thrust	F_T	0 N
lift	0 N	F_L
air resistance	F_R	0 N
weight	$mg = 2.60 \times 10^5$ N	
thrust	$F_T = 3.00 \times 10^5$ N	
flight angle	$\theta = 60.0°$	

What is the strategy?

1. Draw a free-body diagram of the forces on the plane, rotating the axes so three of the forces lie along an axis.
2. Calculate the net force on the plane along the axis of the lift force, and solve for the lift force.

Physics principles and equations

Newton's second law

$$\Sigma\mathbf{F} = m\mathbf{a}$$

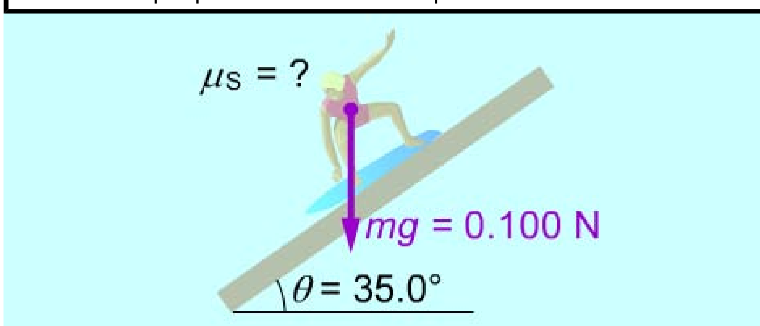
When the acceleration is zero, the forces along every dimension must sum to zero.

Step-by-step solution

The lift force acts in the y dimension. In the y column of the variables table, all the values are known except the lift force. So, we need only apply equilibrium in the y dimension to solve this problem.

Step	Reason
1. $\Sigma\mathbf{F}_y = 0$	no acceleration in y dimension
2. $F_L + (-mg \cos \theta) = 0$	lift and y component of weight
3. $F_L = (2.60 \times 10^5 \text{ N}) \cos 60.0°$	rearrange and enter values
4. $F_L = 1.30 \times 10^5$ N	evaluate

6.7 - Sample problem: an inclined plane and static friction



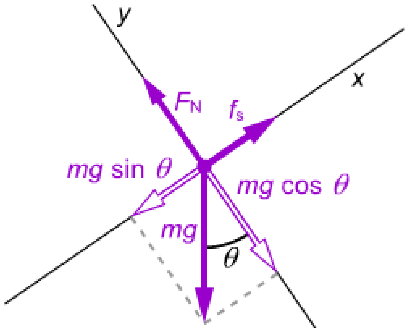
The Surfer Bob toy is just about to slide. What is the coefficient of static friction?

A classic physics lab exercise asks you to use a block on a plane to calculate the coefficient of static friction for two materials. You see that configuration shown above, although instead of a block, we are using Surfer Bob. You are given Bob's weight and the angle the plane makes with the horizontal just before Bob begins to slide. From this information, you are asked to calculate the coefficient of static friction. Since Bob is not accelerating, no net force is acting on him.

You may have performed a lab experiment like this at some point during your studies. You incline a plane until the force of gravity just overcomes friction and causes a block on the plane to slide. You then incline it a little less until the plane is at the angle at which the force of static friction balances the force of gravity down the plane. At this point, the static friction is at its maximum and you can calculate the coefficient.

As you see below, we can solve this problem in fewer steps by rotating the axes. Two of the forces are acting along the inclined plane. By rotating the axes so the x axis is parallel to the plane, we can reduce the amount of trigonometry required. The rotation means that two of the forces act solely along an axis. If we did not do the rotation, each force we analyzed would have to be decomposed into its x and y components with the use of sines and cosines in order for us to solve the problem. If you do not like this axis rotation "trick," then you can always solve the problem by keeping the axes horizontal and vertical and using components.

Draw a free-body diagram



Variables

After we rotate the axes through the angle θ, the component of the force of gravity pulling Bob down the plane, and the frictional force, both lie along the x axis. The component of the weight pressing Bob into the plane and the normal force are on the y axis. Since two forces (the normal force and the frictional force) act solely along an axis, we do not have to calculate their components, and this reduces the amount of work that follows.

When the axes are rotated, correctly identifying the component vectors that result can be difficult. The components are stated below in the variables table. Note that the x component, for instance, is calculated using the sine, not the cosine of the angle θ. The correct trigonometric ratios and signs must be determined using a diagram like the one above.

	x component	y component
weight	$-mg \sin \theta$	$-mg \cos \theta$
normal force	0 N	F_N
friction	f_s	0 N
weight	$mg = 0.100$ N	

What is the strategy?

1. Draw a free-body diagram of the forces acting on Surfer Bob.
2. Bob is not moving, which means he is not accelerating, so the forces in each dimension sum to zero. Sum the forces in the x dimension to get an equation involving the coefficient of static friction.
3. Sum the forces in the y dimension to get an expression for the normal force. Use it to solve the equation previously developed for the coefficient of static friction.

Physics principles and equations

We use the equation for **maximum** static friction

$$f_{s,max} = \mu_s F_N$$

Newton's second law

$$\Sigma \mathbf{F} = m\mathbf{a}$$

When the acceleration is zero, the forces along every dimension must sum to zero.

Step-by-step solution

First we consider the forces in the x dimension. Bob is in equilibrium, so these forces sum to zero.

Step	Reason
1. $\Sigma \mathbf{F}_x = 0$	no acceleration in x dimension
2. $\mu_s F_N + (-mg \sin \theta) = 0$	substitute friction and x component of weight
3. $\mu_s = (mg \sin \theta) / F_N$	solve for μ_s

We now have an equation for the coefficient of static friction, but we do not have enough information to evaluate it. We can, however, use the fact that there is also equilibrium in the y dimension to get more information.

Step	Reason
4. $\Sigma \mathbf{F}_y = 0$	no acceleration in y dimension
5. $F_N + (-mg \cos \theta) = 0$	normal force and y component of weight
6. $F_N = mg \cos \theta$	solve for normal force

Now we substitute the expression for the normal force from step 6 into the equation of step 3 and find the coefficient of static friction.

Step	Reason
7. $\mu_s = (mg \sin\theta) / (mg \cos\theta)$	substitute step 6 into step 3
8. $\mu_s = \sin\theta / \cos\theta$	simplify
9. $\mu_s = \tan\theta$	trigonometric identity
10. $\mu_s = \tan 35.0°$	enter value
11. $\mu_s = 0.700$	evaluate

Even though we were given the weight of the dude, it turns out the coefficient of static friction does not depend on weight, but solely on the angle of the plane.

6.8 - Interactive problem: desert island survival

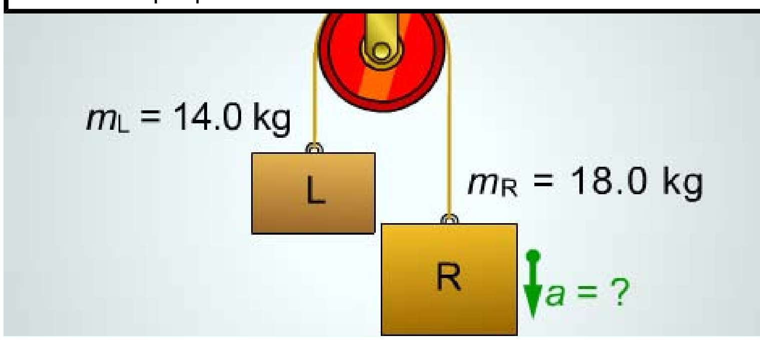
You are stranded on a sweltering desert island. In addition to being sweltering, the island is sloped. Fortunately, you have found a car. What exactly you will do with it remains something of a mystery, but it is better than nothing. Because the car's brakes are broken, you have rigged a pulley system with a hanging block to hold the vehicle stationary under the only tree on the island.

You need to calculate a mass for the hanging block that will keep the car stationary. If the hanging block is too heavy or too light, the car will crash into a rock or roll into the water. The mass of the car is 1210 kg and the angle of the slope from the horizontal is 6.60°. The rope does not stretch and the pulley is massless and frictionless. There is no friction between the car and the ground.

Enter the mass of the block to the nearest kilogram and press GO. If the acceleration is zero, the car will not move, and you will have succeeded. Press RESET to try again. (You may find it more entertaining to fail in this situation, but do try to succeed.)

If you have difficulty solving this, draw a free-body diagram of the forces on the car with the axes rotated so that the x axis matches the slant of the island. The entire force supplied through the rope must then equal a component of the weight.

6.9 - Sample problem: an Atwood machine

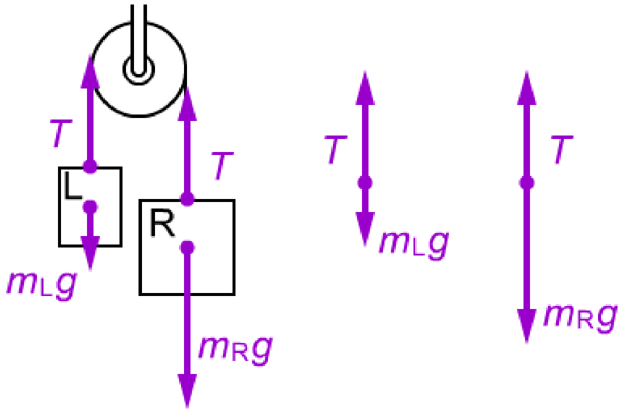


What is the magnitude of acceleration of the blocks?

The illustration above shows two blocks connected by a rope passing over a pulley. Because the blocks partially counterbalance each other, the force required to lift either of them is less than its weight. This type of system is called an *Atwood machine.* An application can be found in elevators, where a massive block partly counterbalances the weight of the elevator car to reduce the force required from the motor that lifts the car.

In this sample problem, you are asked to find the rate at which the blocks accelerate. As usual, the rope and pulley are massless, the rope does not stretch, and the pulley has no friction. The rope exerts an equal tension force on both blocks. The blocks' accelerations are equal in magnitude but opposite in direction.

Draw a free-body diagram



Variables

Block L

tension	T
mass	$m_L = 14.0$ kg
weight	$-m_L g$
acceleration	a

Block R

tension	T
mass	$m_R = 18.0$ kg
weight	$-m_R g$
acceleration	$-a$

The negative sign for the acceleration of block R indicates that it is directed downward. Because we are using the same variable a for the magnitude of the acceleration of the two blocks, we must assign signs to their vertical acceleration components that reflect their directions, as we do in the variables table.

What is the strategy?

1. Draw a free-body diagram for each block.
2. Calculate the net force on each block. There are no horizontal forces on either block.
3. Use Newton's second law for each block to find two expressions for the tension force in the rope, and set the expressions equal to find the acceleration.

Physics principles and equations

Since both blocks move only vertically, we apply Newton's second law in the y dimension.

Step-by-step solution

We start by applying Newton's second law to the block on the left to find an equation for the tension force. All the forces are in the vertical direction, so we only consider the y components.

Step	Reason
1. $\Sigma \mathbf{F}_y = m_L \mathbf{a}_y$	Newton's second law, left block
2. $T + (-m_L g) = m_L a$	net force is tension plus weight
3. $T = m_L g + m_L a$	solve for T

Then we apply Newton's second law to the block on the right, in the vertical direction, and find a second equation for the tension force.

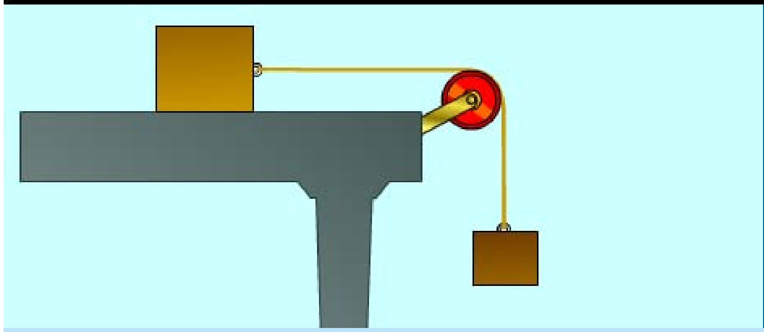
Step	Reason
4. $\Sigma \mathbf{F}_y = m_R \mathbf{a}_y$	Newton's second law, right block
5. $T + (-m_R g) = m_R(-a)$	net force is tension plus weight
6. $T = m_R g - m_R a$	solve for T

We set the two expressions for tension equal and solve for the acceleration.

Step	Reason
7. $m_L g + m_L a = m_R g - m_R a$	set tensions equal, from steps 3 and 6
8. $a = \dfrac{(m_R - m_L)(g)}{m_R + m_L}$	solve for a
9. $a = \dfrac{(18.0 - 14.0\text{ kg})(9.80\text{ m/s}^2)}{18.0\text{ kg} + 14.0\text{ kg}}$	enter values
10. $a = 1.23\text{ m/s}^2$	arithmetic

Equation 8 can be profitably analyzed by considering a couple of special cases. If m_L equals zero, equation 8 states that the acceleration equals g. This makes sense: The block on the right would be in free fall, since no force would be opposing the force of its weight. Also, if $m_R = m_L$, the acceleration would be zero, since there would be no net force. If we let m_R go to infinity, equation 8 states that the acceleration equals g. This too makes sense. If m_R is very much bigger than m_L, then m_L will hardly slow down m_R as it falls in free fall.

6.10 - Interactive checkpoint: two blocks

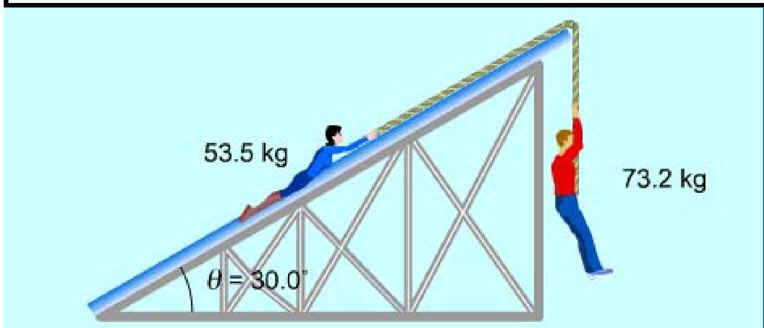


A 15.8 kg block sits on a frictionless horizontal table. The block is attached to a horizontal string that goes over a pulley and is connected to another block that hangs freely. The string is massless and does not stretch. The acceleration of the block on the table is 3.89 m/s^2. What is the mass of the hanging block?

Answer:

m_h = [ ] kg

6.11 - Interactive checkpoint: two "blocks" at an angle

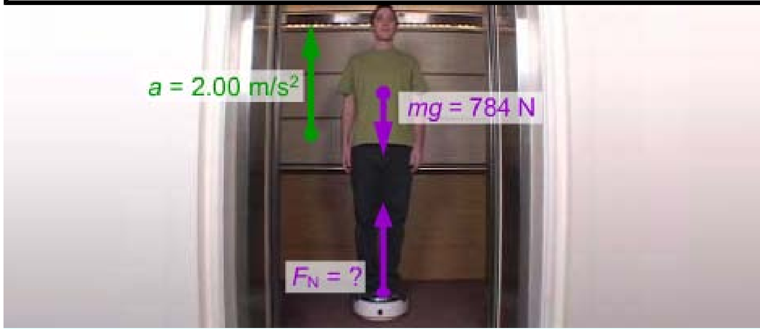


Frances, a 53.5 kg woman, slides on a frictionless, icy ramp that is inclined at 30.0° to the horizontal. An unstretchable rope connects her to Andre, who has a mass of 73.2 kg and is accelerating at the same rate, but parallel to the vertical wall of the ramp. What is the magnitude of their acceleration?

Answer:

a = [ ] m/s^2

6.12 - Sample problem: weight in an accelerating elevator



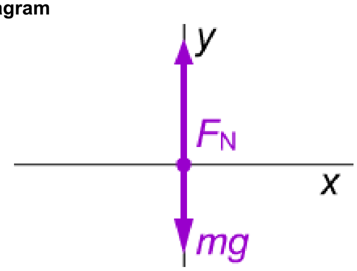
What is the magnitude of the normal force exerted by the scale?

You may be familiar with feeling heavier or lighter in an elevator as it changes speed. When the elevator above is not accelerating, the scale reports the person's *actual weight*. When the person (Kevin) and scale are not accelerating, the scale indicates the amount of the gravitational force that is exerted on Kevin, which equals mg.

When the elevator accelerates upward, Kevin feels heavier and the scale reports a value larger than his actual weight. When they accelerate downward, the scale reports a smaller value. Each of these values is an *apparent weight*. Kevin's actual weight does not change; it is the acceleration that causes the scale to report these apparent weights.

In the problem, you are asked to determine Kevin's apparent weight when the elevator is accelerating upwards at 2.00 m/s. You can do this by determining the normal force that the scale exerts on him. A scale shows the amount of the normal force it is exerting on the person standing on it. The normal force can be determined by applying Newton's second law; the other force acting on Kevin (gravity) and his acceleration are both stated.

Draw a free-body diagram



Variables

actual weight	$-mg = -784$ N
normal force	F_N
acceleration	$a = 2.00$ m/s^2
mass	$m = 784$ N / 9.80 m/s^2 = 80.0 kg

What is the strategy?

1. Draw a free-body diagram of the forces on the person.
2. Determine the net force on the person.
3. Use Newton's second law to find the normal force.

Physics principles and equations

Newton's second law

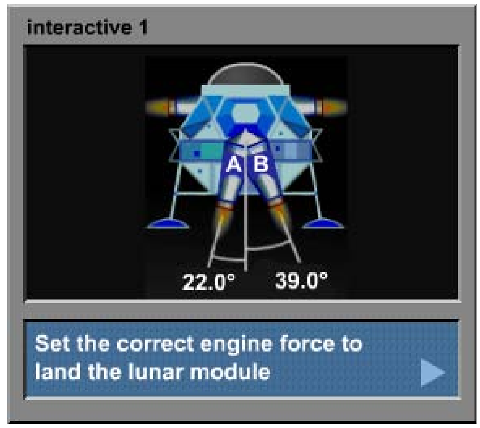
$$\Sigma \mathbf{F} = m\mathbf{a}$$

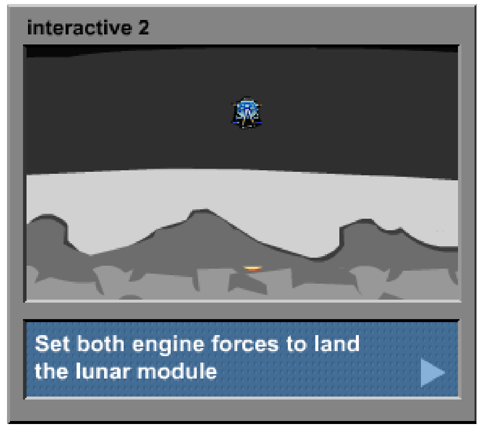
Step-by-step solution

Step	Reason
1. $\Sigma \mathbf{F}_y = m\mathbf{a}_y$	Newton's second law
2. $F_N + (-mg) = ma$	substitute forces
3. $F_N = mg + ma$	solve for normal force
4. $F_N = 784\text{ N} + (80.0\text{ kg})(2.00\text{ m/s}^2)$	enter values
5. $F_N = 944$ N	evaluate

Equation 3 above is worth considering. If the acceleration a is zero, the normal force is the person's true weight, as should be expected. If the elevator is in free fall, the acceleration a equals $-g$, so the terms on the right cancel. That means in free fall the normal force is zero and the scale exerts no force on the person (and vice versa).

6.13 - Interactive summary problem: lunar module landing





You are in training to be an astronaut. In one of your training simulations, your goal is to land a lunar excursion module (LEM) on the surface of the Moon. This LEM has two rocket engines attached to the center of the craft. The engines typically point at angles of $22.0°$ in opposite directions from the negative y axis and they generate equal amounts of thrust. This means that as they fire, they provide a combined upward force that does not cause the rocket to move to the right or left.

In an "accident" intended to test your astronaut skills, the right rocket (labeled B) is bent to a $39.0°$ angle. The illustration to the right shows these two angles. In spite of this problem, you are still expected to guide the LEM straight down.

There are two levels in this exercise. In the first, the module is moving downward toward the landing pad as the simulation starts. In this "emergency," the amount of thrust force from the left engine, labeled A, is jammed at 35,600 N.

You are allowed to set the amount of force from the right engine. You need to set this amount so there is no net horizontal force. If you set engine B's force this way, you will land the LEM gently onto the landing pad. Small attitude rockets at the top of the LEM will keep it from rotating, but it will drift left or right if you enter the wrong values.

To accomplish your first mission, compute the force needed from the right engine, set that value to the nearest 100 N in the simulation, and press GO to test your answer. Press RESET to try again.

The second exercise is harder and is optional. In the second simulation, the LEM is again moving downward toward the landing pad. You need to set the forces of both engines to achieve a net acceleration directly upward of 4.12 m/s^2. The mass of the module is 8910 kg, and it does not change significantly as the engines burn fuel. As an astronaut, you know that the rate of acceleration due to the Moon's gravity is 1.62 m/s^2.

To solve this problem, first calculate the total vertical force that should be provided by the engines to counter the Moon's gravity and provide a net upward acceleration of 4.12 m/s^2. This will give you one equation involving the two unknown engine forces. The horizontal components of the misaligned engine forces must still balance so that the LEM stays on course. This gives you another equation. Solve the two equations and enter the force values to the nearest 100 N, then press GO.

If you have difficulty with this exercise, review the earlier sections in this chapter dealing with forces that do not act solely along one axis.

6.14 - Summary

This chapter focuses on applying concepts in sample problems. The basic concepts are:

1. An object that is stationary has zero acceleration and has no net force acting on it. An object moving at a constant velocity also has no net force acting on it.
2. Draw free-body diagrams. They help you analyze problems.
3. Break forces into components and sum those components along each axis. If the object is not accelerating, the sum of the x components equals zero, as does the sum of the y components.
4. Once you have calculated the net force, use Newton's second law. Divide the net force by the mass to calculate the acceleration. (You may have to divide the weight by the acceleration of gravity to determine the mass.)
5. If there are two unknowns, you will need to develop two equations that include those unknowns.

And practice as needed!

Chapter 6 Problems

Chapter Assumptions

Unless stated otherwise, assume that all pulleys, ropes, strings, wires and other connecting materials are massless, frictionless and otherwise ideal.

Conceptual Problems

C.1 Only one non-zero force acts on an object. Can the object be and remain at rest? Explain.

○ Yes ○ No

C.2 If an object has no velocity, does that necessarily mean no forces are acting on it? Explain your answer.

○ Yes ○ No

C.3 A massless rope is hanging over a massless and frictionless pulley. A large bunch of bananas is tied to one end of the rope, and a monkey of equal weight is clinging to the other end, at a lower height. The monkey decides he wants the bananas and starts to climb his side of the rope. What happens to the bananas as the monkey climbs?

i. The bananas rise at the same rate and stay out of reach
ii. The bananas stay in place
iii. The bananas fall at the same speed at which the monkey climbs

C.4 A car is parked on a downhill section of a hill. What forces are acting on the car?

C.5 Imagine using a spring scale on the Moon, where the acceleration of gravity is less than on Earth. Would the scale display different values on the Moon and Earth? Explain why.

○ Yes ○ No

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) If $\mathbf{F}_1$ is set to 12 N directly to the left, what should $\mathbf{F}_2$ be set to so that the ball does not move when you press GO? (b) If $\mathbf{F}_1$ is set to 10 N directly to the left, what should $\mathbf{F}_2$ be set to so that the ball hits the target directly to the right of the ball? (c) If $\mathbf{F}_1$ is set to 10 N straight up, what should $\mathbf{F}_2$ be set to so that the ball hits the target that is up and to the right of the ball?

(a) __________ N,
i. Straight up
ii. Directly to the right
iii. Straight down
iv. Directly to the left

(b)
i. Any magnitude less than 10 N
ii. Exactly 10 N
iii. Any magnitude greater than 10 N

,
i. Straight up
ii. Directly to the right
iii. Straight down
iv. Directly to the left

(c) __________ N,
i. Straight up
ii. Directly to the right
iii. Straight down
iv. Directly to the left

Section 1 - Sample problem: a mass on ropes

1.1 A child sits still on a swing that is supported by two chains. Each chain makes a 15.0° angle with the vertical. The child's mass is 25.0 kg. What is the tension in each chain? (Ignore the mass of the seat and chains.)

__________ N

1.2 A tightrope walker is practicing. She is balanced on a rope at the exact midpoint of the rope. The tension in the rope is 4000 N, and the mass of the tightrope walker is 51 kg. Consider the left side of the rope. What angle with the horizontal does it make? State this as a positive number in degrees.

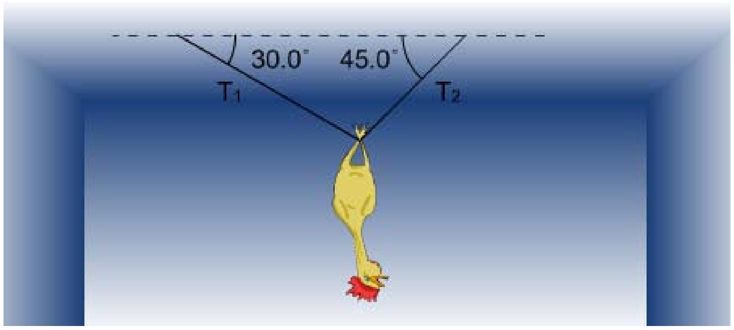
__________ °

1.3 A painting of mass 3.20 kg hangs on a wall. Two thin pieces of wire, each 0.250 m long, connect the painting's center to two hooks in the wall. The hooks are at the same height and are 0.330 m apart. When the painting hangs straight on the wall, how much tension is in each piece of wire?

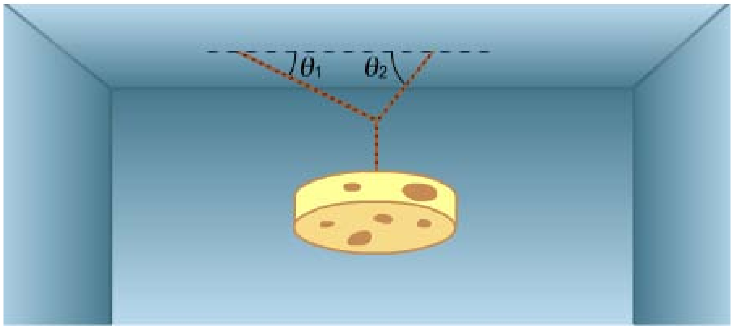
__________ N

1.4 A rubber chicken of mass 0.850 kg is dangling as shown. What is the tension in each string?

T 1 = __________ N
T 2 = __________ N



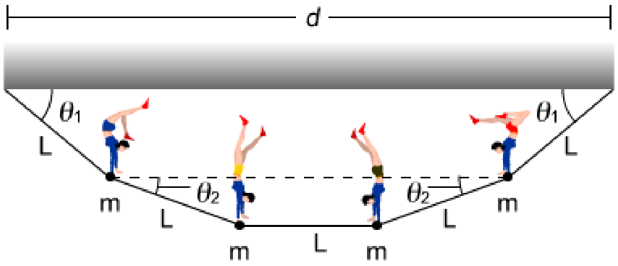
1.5 A 4.50 kg block of gruyere cheese is suspended as shown. Rope 1 on the left has a tension of 28.3 N; rope 2 on the right has a tension of 40.5 N. (a) What is θ_1, the angle made by rope 1? (b) What is θ_2? State each angle as the positive angle "inside" the triangle the ropes form.
You may wish to use at least one of the following trigonometric identities.



$$\arcsin x = \arccos\sqrt{1-x^2}$$
$$\arccos x = \arcsin\sqrt{1-x^2}$$
$$\cos^2\theta + \sin^2\theta = 1$$

(a) __________ °
(b) __________ °

1.6 A circus has identical quadruplet acrobats. Each has a mass *m*. They stand on a tightrope, evenly spaced from one end to the other, each a distance *L* from the next. At each end, the string forms an angle θ_1 with the flat ceiling. The center section of rope is horizontal and parallel to the ceiling. The remaining 2 segments form an angle θ_2 with the horizontal. T_1 is the tension in the leftmost section of tightrope, T_2 is the tension in the section adjacent to it, and T_3 is the tension in the horizontal segment. (a) Find the tension in each section of rope in terms of θ_1, *m*, and *g*. What is the tension T_3? (b) Find the angle θ_2 in terms of θ_1. If θ_1 is 5.10°, what is θ_2? (c) Find the distance *d* between the end points in terms of *L* and θ_1. Note: For all three parts of this question, derive the equations as asked. Parts (a) and (b) ask specific questions so that a computer can evaluate your response; (c) does not.



(a)
- ○ $2mg/\sin\theta_1$
- ○ $mg/\sin\theta_1$
- ○ $2mg/\tan\theta_1$
- ○ *mg*
- ○ None of the above

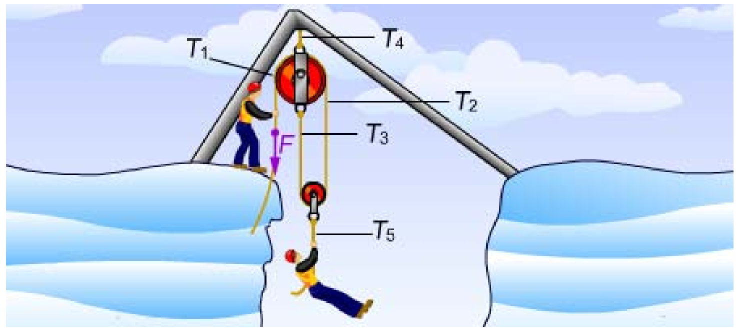
(b) __________ °
(c) Submit answer on paper.

Section 3 - Sample problem: pulling up a scaffold

3.1 A cat is stuck in a tree. You are designated with the job to get it out, yet you do not want to climb the tree, because you may get stuck as well. Instead you set up a pulley system. A rope (consider it massless) runs from the seat you sit on over an ideal pulley and then to your hand. You pull on the loose end of the rope with a force of 348 N. You weigh 612 N and the seat you sit on weighs 16.0 N. (a) What is your acceleration? (b) What force does the seat exert on you?

(a) __________ m/s 2
(b) __________ N

3.2 Your fellow climber is stuck in a crevasse and you need to get him out using a pulley system shown in the diagram. You are pulling your friend, of mass *M*, out of the crevasse at constant speed by applying a force *F*. The pulleys are frictionless and have negligible mass, as do the ropes. Find each tension: (a) T_1, (b) T_2, (c) T_3, (d) T_4, (e) T_5 and (f) the magnitude of *F*. (Hint: Drawing a free body diagram for each pulley will help.)



(a) i. Mg N
ii. Mg/2
iii. 2Mg
iv. 3Mg/2
v. Mg/3
vi. 5Mg
vii. Mg/5

(b) i. Mg N
ii. Mg/2
iii. 2Mg
iv. 3Mg/2
v. Mg/3
vi. 5Mg
vii. Mg/5

(c) i. Mg N
ii. Mg/2
iii. 2Mg
iv. 3Mg/2
v. Mg/3
vi. 5Mg
vii. Mg/5

(d) i. Mg N
ii. Mg/2
iii. 2Mg
iv. 3Mg/2
v. Mg/3
vi. 5Mg
vii. Mg/5

(e) i. Mg N
ii. Mg/2
iii. 2Mg
iv. 3Mg/2
v. Mg/3
vi. 5Mg
vii. Mg/5

(f) i. Mg N
ii. Mg/2
iii. 2Mg
iv. 3Mg/2
v. Mg/3
vi. 5 Mg
vii. Mg/5

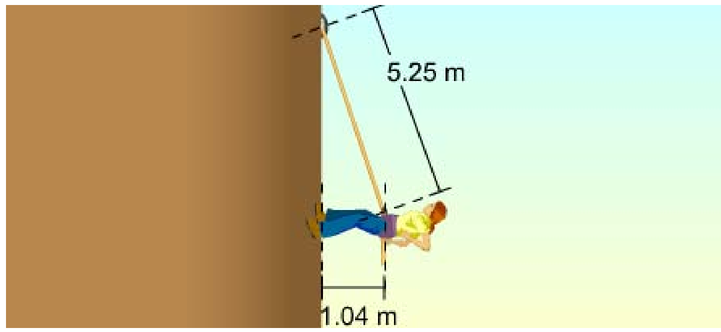
Section 5 - Interactive problem: mountain rescue

5.1 Use the information given in the interactive problem in this section to answer the following question. What is the force required to save the food from the bear? As in the interactive, subtract 2 N from your answer to ensure that the chock does not come loose. Test your answer using the simulation.

________ N

Section 6 - Sample problem: airplane at constant velocity

6.1 A climber of mass 64.8 kg is rappelling down a cliff, but has momentarily paused. She stands with her feet pressed against the icy, frictionless rock face and her body horizontal. A rope of negligible mass is attached to her near her waist, 1.04 m horizontally from the rock face. There is 5.25 m of rope between her waist and where the rope is attached to a chock in the face of the vertical wall she is descending. Calculate the tension in the rope.



________ N

Section 7 - Sample problem: an inclined plane and static friction

7.1 Two people are pushing their stalled car up a hill with an incline of 4.20°. They are pushing parallel to the surface of the hill. The car's mass is 954 kg. What is the combined force they must apply to keep the car moving at a constant speed up the hill? Ignore the forces of friction and air resistance.

________ N

7.2 A block sits on an inclined plane, which is slowly being raised. The block remains motionless until the angle the plane makes with the horizontal is 22°. At this angle, the block begins to slide down the plane. What is the coefficient of static friction between the plane and the block?

____

7.3 You decide to find the coefficient of static friction between a paper cup and an archaic paper textbook. Armed with a metric ruler, you place the cup on the edge of the book, and raise the edge of the book until the cup starts to slide downwards. The length of the book is 26.0 cm, and you need to raise one edge 8.00 cm before the paper cup starts sliding. If the cup has a mass of 211 grams, what is the coefficient of static friction between the cup and the book?

____

7.4 A beaver is pulling a small branch of mass m up a hill at a constant velocity. The hill slope is at an angle θ above the horizontal. The coefficient of friction is μ_k. State how much force he must exert on the branch as he pulls it parallel to the slope in terms of the variables stated in this problem.

- ○ $mg \sin\theta + \mu_k \cos\theta$
- ○ $mg \sin\theta + mg\mu_k \cos\theta$
- ○ $mg \cos\theta + mg\mu_k \sin\theta$
- ○ None of the above

7.5 A skier of mass 64 kg starts from rest at the top of an 18° slope, points her skis straight down the slope and lets gravity pull her down the slope. The slope is 65 m long. Her speed at the bottom is 15 m/s. (a) Assuming zero air resistance, what is the magnitude of the frictional force between her skis and the snow? (b) What is the coefficient of kinetic friction between her skies and the snow?

(a) ________ N
(b) ________

7.6 A block of mass 15.0 kg sits on a plane that is inclined at a 37.0° angle to the horizontal. A 227 N force, pointing up the plane, is applied to the block. The coefficient of kinetic friction is 0.500. What is the speed of the block after 2.00 s?

________ m/s

7.7 When in deep space, a block accelerates at 4.2 m/s^2 when a 120 N force is applied to it. The coefficient of static friction for a particular plane and this block is 0.31. What angle can this plane make with the horizontal before the object starts to slide?

________ °

Section 8 - Interactive problem: desert island survival

8.1 Use the information given in the interactive problem in this section to calculate the mass for the hanging block to keep the car stationary. Test your answer using the simulation.

_________ kg

Section 9 - Sample problem: an Atwood machine

9.1 A 10.4 kg block sits on a frictionless horizontal table. The block is attached to a horizontal ideal string that goes over an ideal pulley and is connected to another identical 10.4 kg block that hangs freely. What is the acceleration of the block on the table? State the acceleration as a positive quantity.

_________ m/s 2

9.2 Two blocks are connected by an ideal string that passes over a massless, frictionless pulley. The two blocks hang freely. The first block has a mass of 14.3 kg, and the second block weighs 98.0 N. Determine (a) the magnitude of the blocks' acceleration, (b) the magnitude of the tension in the string.

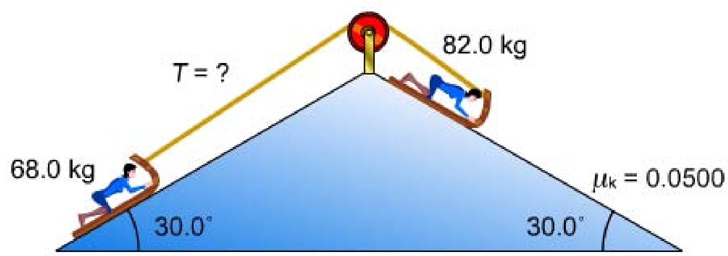
(a) _________ m/s 2
(b) _________ N

9.3 At a construction site, two barrels, each of mass of 122 kg, are connected by a long, massless rope that passes over an ideal pulley that is mounted high overhead. The red barrel is at ground level, and the blue barrel is 8.26 m off the ground, at the third story of a building. Some careless bricklayers hold the blue barrel in place, load it with 200 kg of bricks, and then let it drop. (a) Calculate the maximum speed of the red barrel. (b) What is the maximum height that the red barrel reaches? Remember that it is still moving upward even after the blue barrel hits the ground and stops.

(a) _________ m/s
(b) _________ m

Section 11 - Interactive checkpoint: two “blocks” at an angle

11.1 You and your friend are sledding on two sides of a triangle-shaped hill. On your side, the hill slopes up at 30.0° from the horizontal; on your friend's side, it slopes down at the same angle. You do not want to climb up the hill, so you tell your friend to thread a rope through an ideal pulley that is conveniently atop the hill. He connects the rope to his sled and tosses the other end of the rope to you. The sleds on the snow have a coefficient of kinetic friction, μ_k, of 0.0500. The total mass of your friend and his sled is 82.0 kg while you and your sled have a mass of 68.0 kg. (a) What is the magnitude of the acceleration of each sled? (b) What is the tension in the rope?



(a) _________ m/s 2
(b) _________ N

Section 13 - Interactive summary problem: lunar module landing

13.1 Use the information given in the first interactive problem in this section to calculate the force required from the right engine to land the lunar module. Test your answer using the simulation.

_________ N

13.2 Use the information given in the second interactive problem in this section to calculate the forces required by (a) the right engine and (b) the left engine to land the lunar module safely. Test your answer using the simulation.

(a) _________ N
(b) _________ N

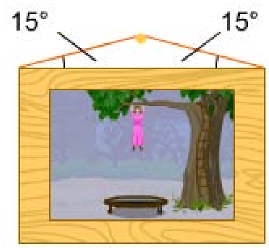
Additional Problems

A.1 You apply a force at an angle θ below the horizontal on a free-sliding sofa. Determine an expression for the acceleration of the sofa as a function of the force *F* you apply, the sofa's mass *m*, and θ.

- ( ) $(F \sin \theta)/m$
- ( ) $Fm \sin \theta$
- ( ) $F \sin \theta$
- ( ) $(F \cos \theta)/m$

A.2 A wire of negligible mass is used to hang a 2.15 kg picture. The ends of the wire are attached to the top 2 corners of the picture. Each end of the wire forms an angle of 15.0 degrees from the horizontal. What is the tension in the wire?

__________ N



A.3 A llama pulls on a sled with a force given by **F** = 2.54**i** + 5.26**j**. A horse pulls on the same sled with a force defined by **F** = 4.25**i** − 6.12**j**. The friction-free sled has a mass of 325 kg, and the forces are in newtons. (a) What is the net force on the sled? (b) What is the acceleration of the sled? (c) If the sled starts at rest, what is its velocity after 4.50 seconds?

(a) __________ i + __________ j N
(b) __________ i + __________ j m/s 2
(c) __________ i + __________ j m/s

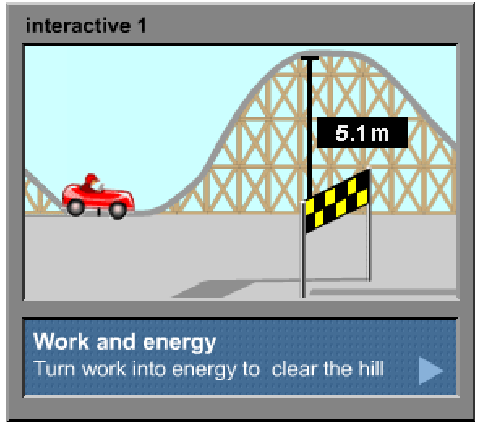
A.4 A 0.0820 kg pair of fuzzy dice is attached to the rearview mirror of a car by a short string. The car accelerates at a constant rate, and the fuzzy dice hang at an angle due to the car's acceleration. If the dice hang at 5.50° from the vertical as the car accelerates, what is the acceleration of the car?

__________ m/s 2

A.5 A child pulls on a toy locomotive of mass 0.979 kg with a force of 3.25 N to the right. The locomotive is connected to two train cars by cables. Friction in the axles results in an effective coefficient of kinetic friction between the floor and the train which is 0.110. One car has a mass of 0.952 kg and the other has a mass of 0.419 kg. (a) What is the acceleration of the train? (b) What is the tension in the cable between the locomotive and the car connected to the locomotive?

(a) __________ m/s 2
(b) __________ N

7.0 - Introduction



The use of energy has played an important role in defining much of human history. Fire warmed and protected our ancestors. Coal powered the Industrial Revolution. Gasoline enabled the proliferation of the automobile, and electricity led to indoor lighting, then radio, television and the computer. The enormous energy unleashed by splitting the atom was a major factor in ending World War II. Today, businesses involved in technology or media may garner more newspaper headlines, but energy is a larger industry.

Humankind has long studied how to harness and transform energy. Early machines used the energy of flowing water to set wheels spinning to mill grain. Machines designed during the Industrial Revolution used energy unleashed by burning coal to create the steam that drove textile looms and locomotives. Today, we still use these same energy sources − water and coal − but often we transform the energy into electric energy.

Scientists continue to study energy sources and ways to store energy. Today, environmental concerns have led to increased research in areas including atomic fusion and hydrogen fuel cells. Even as scientists are working to develop new energy technologies, there is renewed interest in some ancient energy sources: the Sun and the wind. They too can provide clean energy via photovoltaic cells and wind turbines.

Why is energy so important? Because humankind uses it to do work. It no longer requires as much human labor to plow fields, to travel, or to entertain ourselves. We can tap into other energy sources to serve those needs.

This chapter is an introduction to work and energy. It appears in the mechanics section of the textbook, because we focus here on what is called *mechanical energy*, energy arising from the motion of particles and objects, and energy due to the force of gravity. Work and energy also are major topics in thermodynamics, a topic covered later. Thermodynamics adds the topic of heat to the discussion. We will only mention heat briefly in this chapter.

Whatever the source and ultimate use of energy, certain fundamental principles always apply. This chapter begins your study of those principles, and the simulation to the right is your first opportunity to experiment with them.

Your mission in the simulation is to get the car over the hill on the right and around a curve that is beyond the hill. You do this by dragging the car up the hill on the left and releasing it. If you do not drag it high enough, it will fail to make it over the hill. If you drag it too high, it will fly off the curve after the hill. The height of the car is shown in a gauge in the simulation.

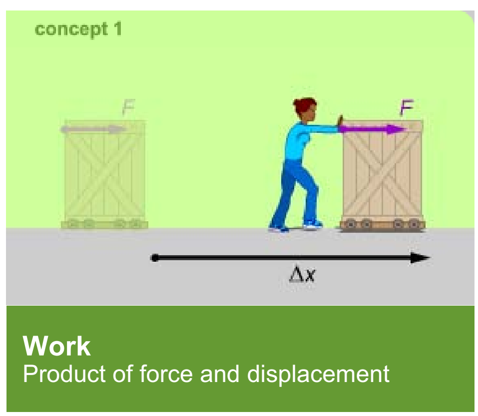
Only the force of gravity is factored into this simulation; the forces of friction and air resistance are ignored. In this chapter, we consider only the kinetic energy due to the object moving as a whole and ignore rotational energy, such as the energy of the car wheels due to their rotational motion. (Taxes, title and dealer prep are also not factored into the simulation; contact your local dealership for any other additional restrictions or limitations.)

Make some predictions before you try the simulation. If you release the car at a higher point, will its speed at the bottom of the hill be greater, the same or less? How high will you have to drag the car to have it just reach the summit of the other hill: to the same height, higher or lower? You can use PAUSE to see the car's speed more readily at any point.

When you use this simulation, you are experimenting with some of the key principles of this chapter. You are doing work on the car as you drag it up the hill, and that increases the car's energy. That energy, called potential energy, is transformed into kinetic energy as the car moves down the hill. Energy is conserved as the car moves down the hill. It may change forms from potential energy to kinetic energy, but as the car moves on the track, its total energy remains constant.

7.1 - Work

Work: The product of displacement and the force in the direction of displacement.



You may think of work as homework, or as labor done to earn money, or as exercise in a demanding workout.

But physicists have a different definition of work. To them, work equals the component of force exerted on an object along the direction of the object's displacement, times the object's displacement.

When the force on an object is in the same direction as the displacement, the magnitude of the force and the object's displacement can be multiplied together to calculate the work done by the force. In Concept 1, a woman is shown pushing a crate so that all her force is applied in the same direction as the crate's motion.

All the force need not be in the direction of the displacement. When the force and displacement vectors are not in the same direction, only the component of the force in the direction of the displacement contributes to work. Consider the woman pulling the crate at an angle with a handle, as shown in Concept 2. Again, the crate slides along the ground. The component of the force perpendicular to the displacement contributes nothing to the work because there is no motion up or down.

Perhaps subconsciously, you may have applied this concept. When you push on a heavy object that is low to the floor, like a sofa, it is difficult to slide it if you are mostly pushing down on top of it. Instead, you bend low so that more of your force is horizontal, parallel to the desired motion.

The equation on the right is used to calculate how much work is done by a force. It has notation that may be new to you. The equation states that the work done equals the "dot product" of the force and displacement vectors (the name comes from the dot between the $\mathbf{F}$ and the $\Delta\mathbf{x}$).

The equation is also expressed in a fashion that you will find useful: $(F \cos \theta)\Delta x$. The angle θ is the angle between the force and displacement vectors when they are placed tail to tail. The vectors and the angle are shown in Equation 1.

By multiplying the amount of force by cos θ, you calculate the component of the force parallel to the displacement. You may recall other cases in which you used the cosine or sine of an angle to calculate a component of a vector. In this case, you are calculating the component of one vector that is parallel to another.

The equation to the right is for a constant or average force. If the force varies as the motion occurs, then you have to break the motion into smaller intervals within which the force is constant in order to calculate the total work.

The everyday use of the word "work" can lead you astray. In physics, if there is no displacement, there is no work. Suppose the woman on the right huffed and puffed and pushed the crate as hard as she could for ten minutes, but it did not move. She would certainly believe she had done work. She would be exhausted. But a physicist would say she has done zero work on the crate because it did not move. No displacement means no work, regardless of how much force is exerted.

Work can be a positive or negative value. Positive work occurs when the force and displacement vectors point in the same direction. Negative work occurs when the force and displacement vectors point in opposite directions. If you kick a stationary soccer ball, propelling it downfield, you have done positive work on the ball because the force and the displacement are in the same direction.

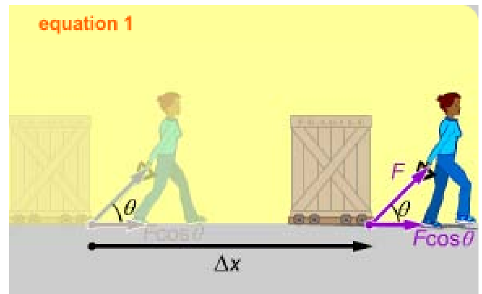
When a goalie catches a kicked ball, negative work is done by the force from the goalie's hands on the ball. The force on the ball is in the opposite direction of the ball's displacement, with the result that the ball slows down.

The sign of work can be calculated with the equation to the right. When force and displacement point in the same direction, the angle between them is $0°$, and the cosine of $0°$ is positive one. When force and displacement point in opposite directions, the angle between the vectors is $180°$, and the cosine of $180°$ is negative one. This mathematically confirms the points made above: Force in the direction of motion results in positive work; force opposing the motion results in negative work.

Work is a scalar quantity, which means it has magnitude but no direction. The *joule* is the unit for work. The units that make up the joule are kg·m^2/s^2 and come from multiplying the unit for force (kg·m/s^2) by the unit for displacement (m).

If several forces act on an object, each of them can do work on the object. You can calculate the *net work* done on the object by all the forces by calculating the net force and using the equation in Equation 1.

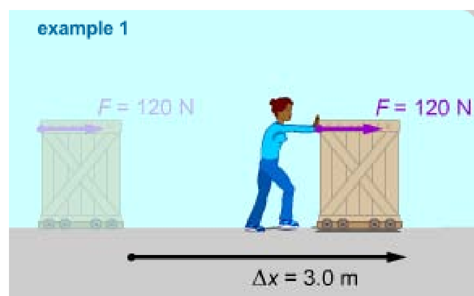
Force at angle to displacement
Only force component along displacement contributes to work



Force at angle to displacement

$$W = \mathbf{F} \cdot \Delta\mathbf{x} = (F \cos \theta)\Delta x$$

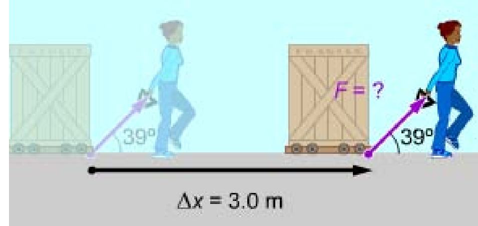
W = work
F = force
Δx = displacement
θ = angle between force and displacement
Unit: joule (J)



How much work does the woman do on the crate?

$W = (F \cos \theta)\Delta x$
$W = (120 \text{ N})(\cos 0°)(3.0 \text{ m})$
$W = (120 \text{ N})(1)(3.0 \text{ m}) = 360 \text{ J}$

example 2



Now the woman is pulling the crate at an angle. If she does the same amount of work as before, how much force must the woman exert?

$W = (F \cos \theta)\Delta x$

$F = W / (\cos \theta)\Delta x$

$F = (360\text{ J}) / (\cos 39°)(3.0\text{ m})$

$F = 150\text{ N}$

7.2 - Dot product

Dot product: The product of the magnitudes of two vectors and the cosine of the angle between the vectors. The result is a scalar.

There are two distinct ways to multiply vectors. In this section we explain in depth one way, called the dot product. The other way to multiply vectors is called the cross product, which is a topic we cover when it is needed later.

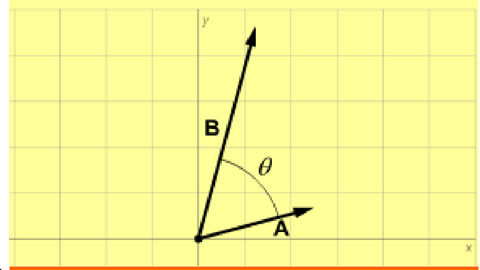
On the right, you see the equation for a dot product. The dot product is more formally called the *scalar product*, but we (and most others) use the term dot product because of the "dot" in the notation. The result of a dot product multiplication is a scalar.

The dot product equals the product of the magnitudes of the vectors and the cosine of the angle between the vectors. This angle is shown in the diagram in Equation 1. To determine the angle, the two vectors are placed tail to tail.

You can think of this multiplication operation as measuring how much of one vector lies along the direction of the other. When the vectors are parallel, the value is maximized, because the angle between the vectors is zero, and $\cos 0°$ equals positive one. When the vectors are perpendicular, the result is zero, because $\cos 90°$ equals zero. At $90°$, no component of a vector lies along the other. When they point in opposite directions, the cos $180°$ equals negative one, and the result is the maximum negative value.

Sometimes you want to calculate a dot product when you know the components of two vectors, but not the angle between them. In this case the second formula in Equation 1 provides a way to calculate the dot product. The components of the two vectors are multiplied as shown, and the values are summed to calculate the result. Because the dot product is commutative, the order of vectors in the operation does not matter. The formula is given for three-dimensional vectors, but with two-dimensional vectors you can use the formula by omitting the a_3 and b_3 terms.

equation 1



Dot product: multiplying vectors

$\mathbf{A} \cdot \mathbf{B} = AB \cos \theta$

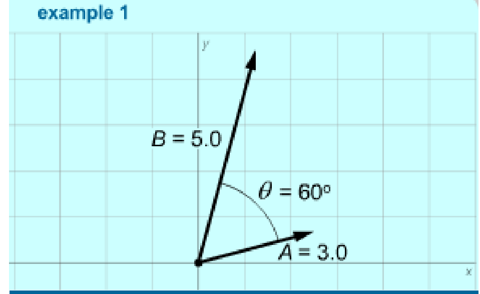
$\mathbf{A} \cdot \mathbf{B}$ = (product of magnitudes)(cosine of angle)

For components:

$\mathbf{A} = (a_1, a_2, a_3)$, $\mathbf{B} = (b_1, b_2, b_3)$

$\mathbf{A} \cdot \mathbf{B} = a_1 b_1 + a_2 b_2 + a_3 b_3$

example 1



What is the dot product of these vectors?

$\mathbf{A} \cdot \mathbf{B} = AB \cos \theta$

$\mathbf{A \cdot B} = (3.0)(5.0)\cos(60°)$
$\mathbf{A \cdot B} = (15)(0.5)$
$\mathbf{A \cdot B} = 7.5$

7.3 - Work done by a variable force

A variable force is a force whose direction or magnitude changes over time. In this section, we discuss a one-dimensional force. Its magnitude can change, as well as its direction along a line.

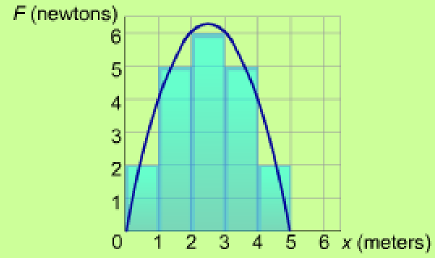
In the graph in Concept 1, the force on an object is plotted on the vertical axis and the displacement of the object is on the horizontal axis. The force is being applied in the direction of the displacement. The graph shows that the force is at its maximum when the displacement is 2.5 meters. The force then decreases until it reaches zero at a displacement of 5.0 meters.

We chop the displacement of the object into smaller segments, as shown in the graph. Using the work equation $W = F \cdot \Delta x$, we find that the work done on the object over a particular segment is approximately equal to the displacement of that segment multiplied by the average force applied during that displacement. The graph shows that the width of a segment multiplied by the average force over that segment is **also** equal to the area of a rectangle in the diagram. Summing the areas of all the rectangles is then approximately equal to the work done over the total displacement.

Using narrower segments would give a closer approximation of the work, because the force for each segment would be averaged over a smaller displacement. As the segments are made increasingly narrow, the total area of the rectangles approaches the exact value of the work, and it also approaches the area under the force-versus-displacement curve. This leads us to conclude that the amount of work done by a variable force equals the area under the force-versus-displacement curve. To be precise, by "area under the curve", we mean the area of the region between the curve and the x axis. The work is positive for an area above the x axis, and negative for an area below.

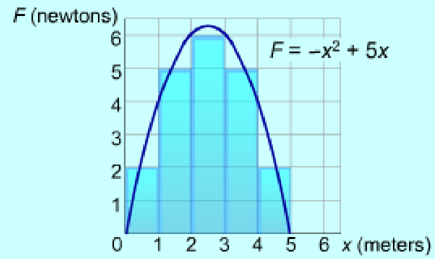
The example problem to the right demonstrates how to estimate the work done by a variable force. We fitted five rectangles to the curve and summed their areas. Using five rectangles yields an answer of 20 J for the work performed by the variable force. We used calculus to determine an exact answer: 20.8 J. Although a larger number of thinner rectangles would make for a more accurate approximation, the approximation shown here is within 4% of the true value.

concept 1



Work by a variable force
Equals area of force, displacement curve
Rectangles approximate area

example 1

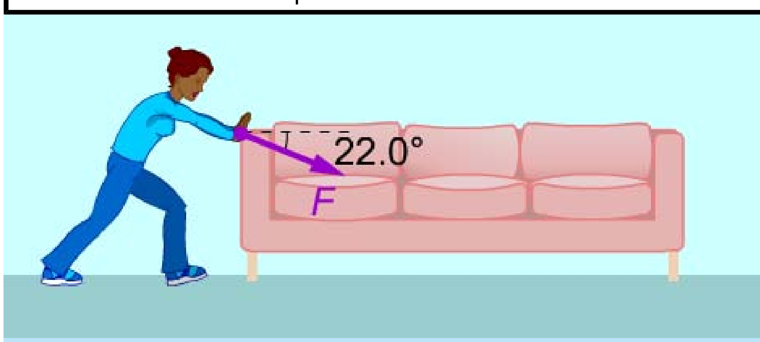


Approximate the work done by the force as it moves the object from 0 to 5 meters.

$W = 2 + 5 + 6 + 5 + 2$

$W = 20\text{ J}$

7.4 - Interactive checkpoint: work



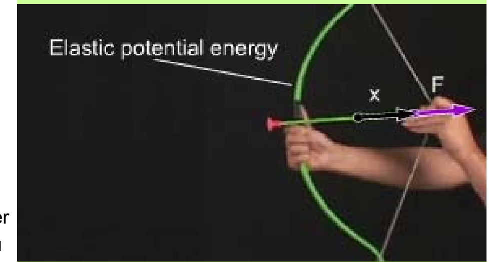
Sally does 401 J of work moving a couch 1.30 meters. If she applies a constant force at an angle of 22.0° to the horizontal as shown, what is the magnitude of this force?

Answer:

$F =$ ______ N

7.5 - Energy

concept 1



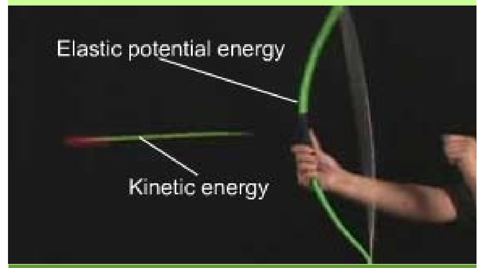
Energy
Is changed by work

Before delving into some specific forms of energy, in this section we address the general topic of energy. Although it is a very important concept in physics, and an important topic in general, energy is notoriously hard to define.

Why? There are several reasons. Many forms of energy exist: electric, atomic, chemical, kinetic, potential, and so on. Finding a definition that fits all these forms is challenging. You may associate energy with motion, but not all forms of energy involve motion. A very important class of energy, potential energy, is based on the position or configuration of objects, not their motion.

Energy is a property of an object, or of a system of objects. However, unlike many other properties covered so far in this textbook, is hard to observe and measure directly. You can measure most forces, such as the force of a spring. You can see speed and decide which of two objects is moving faster. You can use a stopwatch to measure time. Quantifying energy is more elusive, because energy depends on multiple factors, such as an object's mass and the square of its speed, or the mass and positions of a system of objects.

concept 2



Energy
Transfers between objects
Exists in many forms

Despite these caveats, there are important principles that concern all forms of energy. First, there is a relationship between work and energy. For instance, if you do work by kicking a stationary soccer ball, you increase a form of its energy called kinetic energy, the energy of motion.

Second, energy can transfer between objects. When a cue ball in the game of pool strikes another ball, the cue ball slows or stops, and the other ball begins to roll. The cue ball's loss of energy is the other ball's gain.

Third, energy can change forms. When water falls over a dam, its energy of position becomes the energy of motion (kinetic energy). The kinetic energy from the moving water can cause a turbine to spin in a dam, generating electric energy. If that electricity is used to power a blender to make a milkshake, the energy is transformed again, this time into the rotational kinetic energy of the blender's spinning blades.

In Concept 1, an archer does work by applying a force to pull a bowstring. This work increases the elastic potential energy of the bow. When the string is released, it accelerates the arrow, transferring and transforming the bow's elastic potential energy into the kinetic energy of the arrow.

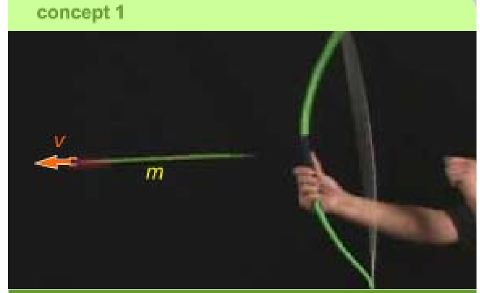
One can trace the history of the energy in the bow and arrow example much farther back. Maybe the chemical energy in the archer that was used by his muscles to stretch the bow came from the chemical energy of a hamburger, and the cow acquired that energy by digesting plants, which got energy via photosynthesis by tapping electromagnetic energy, which came from nuclear reactions in the Sun. We could go on, but you get the idea.

Energy is a scalar. Objects can have more or less energy, and some forms of energy can be positive or negative, but energy does not have a direction, only a value. The joule is the unit for energy, just as it is for work. The fact that work and energy share the same unit is another indication that a fundamental relationship exists between them.

7.6 - Kinetic energy

Kinetic energy: The energy of motion.

concept 1



Kinetic energy
Energy of motion
Proportional to mass, square of speed

Physicists describe the energy of objects in motion using the concept of kinetic energy (KE). Kinetic energy equals one-half an object's mass times the square of its speed.

To the right is an arrow in motion. The archer has released the bowstring, causing the arrow to fly forward. A fundamental property of the arrow changes when it goes from motionless to moving: It gains kinetic energy.

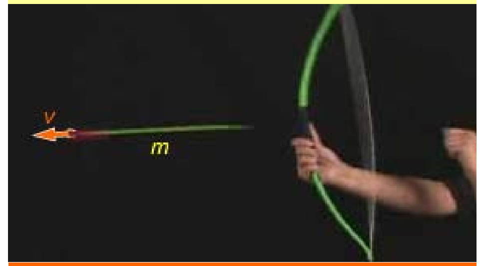
The kinetic energy of an object increases with mass and the square of speed. A 74,000 kg locomotive barreling along at 40 m/s has four times as much kinetic energy as when it is going 20 m/s, and about five million times the kinetic energy of a 6-kg bowling ball rolling at 2 m/s.

With kinetic energy, only the magnitude of the velocity (the speed) matters, not direction. The locomotive, whether heading east or west, north or south, has the same kinetic energy.

Objects never have negative kinetic energy, only zero or positive kinetic energy. Why? Kinetic energy is a function of the speed squared and the square of a value is never negative.

Because it is a type of energy, the unit for kinetic energy is the joule, which is one kg·(m/s)^2. This is the product of the units for mass and the square of the units for velocity.

equation 1

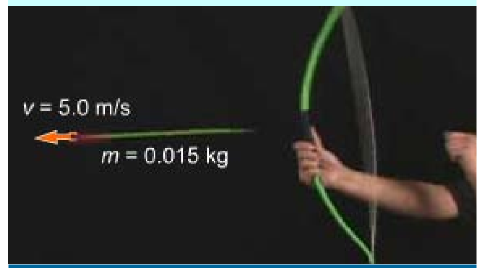


Kinetic energy

$KE = \frac{1}{2} mv^2$

KE = kinetic energy
m = mass
v = speed
Unit: joule (J)

example 1



What is the kinetic energy of the arrow?

$KE = \frac{1}{2} mv^2$

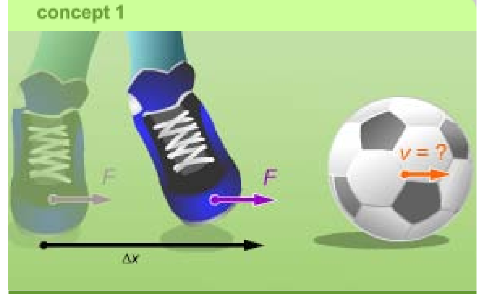
$KE = \frac{1}{2}(0.015 \text{ kg})(5.0 \text{ m/s})^2$

$KE = 0.19$ J

7.7 - Work-kinetic energy theorem

Work-kinetic energy theorem: The net work done on a particle equals its change in kinetic energy.

concept 1



Work done on a particle
Net work equals change in kinetic energy
Positive work on object increases its *KE*

Consider the foot kicking the soccer ball in Concept 1. We want to relate the work done by the force exerted by the foot on the ball to the ball's change in kinetic energy. To focus solely on the work done by the foot, we ignore other forces acting on the ball, such as friction.

Initially, the ball is stationary. It has zero kinetic energy because it has zero speed. The foot applies a force to the ball as it moves through a short displacement. This force accelerates the ball. The ball now has a speed greater than zero, which means it has kinetic energy. The work-kinetic energy theorem states that the work done by the foot on the ball equals the change in the ball's kinetic energy. In this example, the work is positive (the force is in the direction of the displacement) so the work increases the kinetic energy of the ball.

As shown in Concept 2, a goalie catches a ball kicked directly at her. The goalie's hands apply a force to the ball, slowing it. The force on the ball is opposite the ball's displacement, which means the work is negative. The negative work done on the ball slows and then stops it, reducing its kinetic energy to zero. Again, the work equals the change in energy; in this case, negative work on the ball decreases its energy.

In the scenarios described here, the ball is the object to which a force is applied. But you can also think of the soccer ball doing work. The ball applies a force on the goalie, causing the goalie's hands to move backward. The ball does positive work on the goalie because the force it applies is in the direction of the displacement of the goalie's hands.

When stated precisely, which is always worthwhile, the work-kinetic energy theorem is defined to apply to a particle: The net work done on a particle equals the change in its KE. A *particle* is a small, indivisible point of mass that does not rotate, deform, and so on. Various properties of the particle can be observed at any point in time, and by recording those properties at any instant, its *state* can be defined.

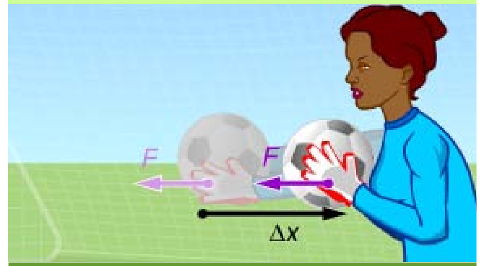
The only form of energy that a single particle can possess is kinetic energy. Stating the work-kinetic energy theorem for a particle means that the work contributes solely to the change in the particle's kinetic energy. A soccer ball is **not** a particle. When you kick a soccer ball, the surface of the ball deforms, the air particles inside move faster, and so forth.

Having said this, we (and others) apply the work-kinetic energy theorem to objects such as soccer balls. A textbook filled solely with particles would be a drab textbook indeed. We simplify the situation, modeling the ball as a particle, so that we can apply the work-kinetic energy theorem. We can always make it more complicated (have the ball rotate or lift off the ground, so rotational KE and gravitational potential energy become factors), but the work-kinetic energy theorem provides an essential starting point.

For instance, in Example 1, we first calculate the work done on the ball by the foot. We then use the work-kinetic energy theorem to equate the work to the change in KE of the ball. Using the definition of KE, we can calculate the ball's speed immediately after being kicked.

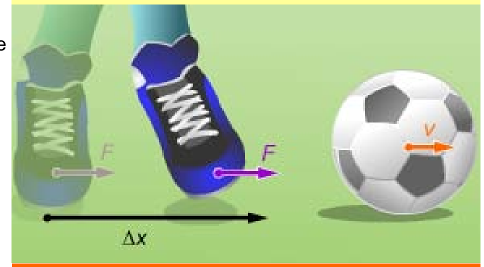
It is important that the theorem applies to the **net** work done on an object. Here, we ignore the force of friction, but if it were being considered, we would have to first calculate the net force being applied on the ball in order to consider the net work that is done on it.

concept 2



Negative work on object
Decreases object's kinetic energy

equation 1

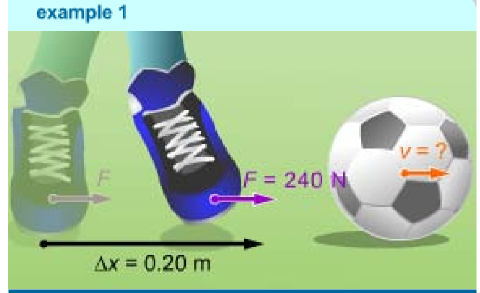


Work-kinetic energy theorem

$$W = \Delta KE$$

W = net work
KE = kinetic energy

example 1



What is the soccer ball's speed immediately after being kicked? Its mass is 0.42 kg.

$W = \mathbf{F} \cdot \Delta\mathbf{x}$

$W = (240\text{ N})(0.20\text{ m}) = 48\text{ J}$

$W = \Delta KE = 48\text{ J}$

$KE = ½\, mv^2 = 48\text{ J}$

$v^2 = 2(48\text{ J})/0.42\text{ kg}$

$v = 15\text{ m/s}$

7.8 - Derivation: work-kinetic energy theorem

In this section, we show that the net work done on an object and its change in kinetic energy are equal by using the definition of work and Newton's second law.

We will again use the illustration of a soccer ball being kicked and model the ball as a particle. The ball starts at rest and we assume the force applied by the foot equals the net force on the ball, and that the ball moves without rotating.

Variables

work	W
force	F
displacement of object	Δx
mass of object	m
acceleration of object	a
initial speed of object	v_i
final speed of object	v
kinetic energy	KE

equation 1

Work and kinetic energy

$W = \Delta KE$

W = work
KE = kinetic energy

Strategy

1. Start with the definition of work.
2. Use Newton's second law to replace the net force in the definition of work by mass times the acceleration.
3. Use a motion equation from the study of kinematics to replace acceleration times displacement with one-half the speed squared.

Physics principles and equations

We will use the definition of work for when the force is in the direction of displacement.

$$W = F\Delta x$$

Newton's second law

$$F = ma$$

Linear motion equation

$$v^2 = v_i^2 + 2a\Delta x$$

The definition of kinetic energy

$$KE = ½mv^2$$

Step-by-step derivation

State the definition of work and use Newton's second law to substitute ma for force.

Step	Reason
1. $W = F\Delta x$	definition of work
2. $W = ma\Delta x$	Newton's second law

We need to replace the acceleration and displacement terms with speed squared to end up with the definition of kinetic energy. We use a motion equation to make this substitution.

Step	Reason
3. $v^2 = v_i^2 + 2a\Delta x$	motion equation
4. $a\Delta x = ½\, v^2$	set $v_i = 0$ and rearrange
5. $W = ½\, mv^2$	substitute equation 4 into equation 2
6. $W = KE$	definition of kinetic energy
7. $W = \Delta KE$	no initial kinetic energy

7.9 - Sample problem: work-kinetic energy theorem

Four bobsledders push their 235 kg sled with a constant force, moving it from rest to a speed of 10.0 m/s along a flat, 50.0-meter-long icy track. Ignoring friction and air resistance, what force does the team exert on the sled?

We assume here that all the work done by the athletes goes to increasing the kinetic energy of the sled.

Variables

mass of sled	$m = 235$ kg
displacement	$\Delta x = 50.0$ m
sled's initial speed	$v_i = 0$ m/s
sled's final speed	$v_f = 10.0$ m/s
work	W
force	F

What is the strategy?

1. Calculate the change in kinetic energy of the sled, using the sled's mass and its initial and final speeds.
2. Use the work-kinetic energy theorem and the definition of work to find the force exerted on the sled.

Physics principles and equations

The definition of work, applied when the force is in the direction of the displacement

$$W = F\Delta x$$

The definition of kinetic energy

$$KE = 1/2\ mv^2$$

The work-kinetic energy theorem

$$W = \Delta KE$$

Step-by-step solution

Start by calculating the change in kinetic energy of the sled.

Step	Reason
1. $\Delta KE = KE_f - KE_i$	definition of change in kinetic energy
2. $\Delta KE = \frac{1}{2} mv_f^2 - \frac{1}{2} mv_i^2$	definition of kinetic energy
3. $\Delta KE = \frac{1}{2} m(v_f^2 - v_i^2)$	factor
4. $\Delta KE = \frac{1}{2}(235\text{ kg})((10.0\text{ m/s})^2 - (0\text{ m/s})^2)$	enter values
5. $\Delta KE = 11{,}800$ J	solve

Use the work-kinetic energy theorem to find the work done on the sled. Then, use the definition of work to determine how much force was exerted on the sled.

Step	Reason
6. $W = \Delta KE$	work-kinetic energy theorem
7. $W = (F \cos \theta)\Delta x$	definition of work
8. $(F \cos \theta)\Delta x = \Delta KE$	set two work equations equal
9. $F \cos \theta = \Delta KE/\Delta x$	rearrange
10. $F = \Delta KE/\Delta x$	force in direction of displacement
11. $F = 11{,}800\ \text{J}/50.0\ \text{m}$	enter values
12. $F = 236\ \text{N}$	solve

7.10 - Interactive problem: work-kinetic energy theorem

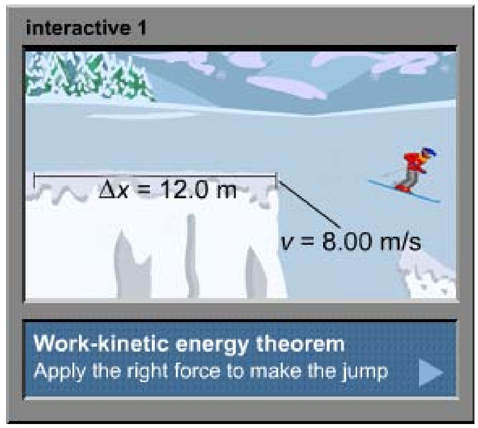
In this simulation, you are a skier and your challenge is to do the correct amount of work to build up enough energy to soar over the canyon and land near the lip of the slope on the right.

You, a 50.0 kg skier, have a flat 12.0 meter long runway leading up to the lip of the canyon. In that stretch, you must apply a force such that at the end of the straightaway, you are traveling with a speed of 8.00 m/s. Any slower, and your jump will fall short. Any faster, and you will overshoot.

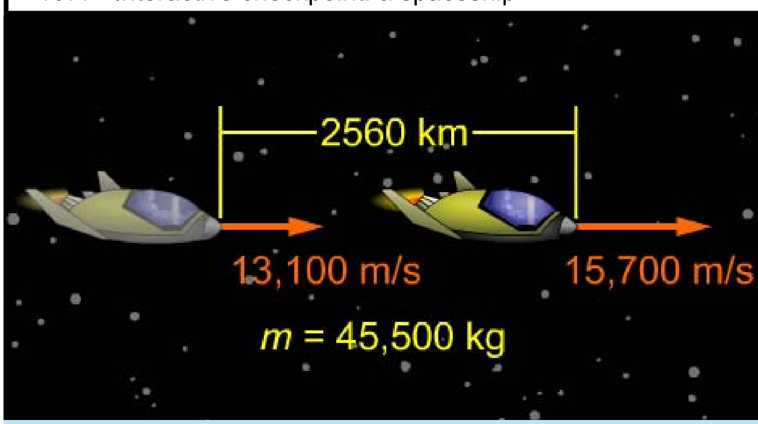
How much force must you apply, in newtons, over the 12.0 meter flat stretch? Ignore other forces like friction and air resistance.

Enter the force, to the nearest newton, in the entry box and press GO to check your result.

If you have trouble with this problem, review the section on the work-kinetic energy theorem. (If you want to, you can check your answer using a linear motion equation and Newton's second law.)



7.11 - Interactive checkpoint: a spaceship



A 45,500 kg spaceship is far from any significant source of gravity. It accelerates at a constant rate from 13,100 m/s to 15,700 m/s over a distance of 2560 km. What is the magnitude of the force on the ship due to the action of its engines? Use equations involving work and energy to solve the problem, and assume that the mass is constant.

Answer:

$F =$ ______ N

Power: Work divided by time; also the rate of energy output or consumption.

The definition of work − the dot product of force and displacement − does not say anything about how long it takes for the work to occur. It might take a second, or a year, or any interval of time. Power adds the concept of time to the topics of work and energy. Power equals the amount of work divided by the time it took to do the work. You see this expressed as an equation in Equation 1.

One reason we care about power is because more power means that work can be accomplished faster. Would you rather have a car that accelerated you from zero to 100 kilometers per hour in five seconds, or five minutes? (Some of us have owned cars of the latter type.)

The unit of power is the watt (W), which equals one joule per second. It is a scalar unit. Power can also be expressed as the rate of change of energy. For instance, a 100 megawatt power plant supplies 100 million joules of energy to the electric grid every second.

Sometimes power is expressed in terms of an older unit, the horsepower. This unit comes from the days when scientists sought to establish a standard for how much work a horse could do in a set amount of time. They then compared the power of early engines to the power of a horse. One horsepower equals 550 foot-pounds/second, which is the same as 746 watts.

We still measure the power of cars in horsepower. For instance, a 300-horsepower Porsche is more powerful than a 135-horsepower Toyota. The Porsche's engine is capable of doing more work in a given period of time than the Toyota's.

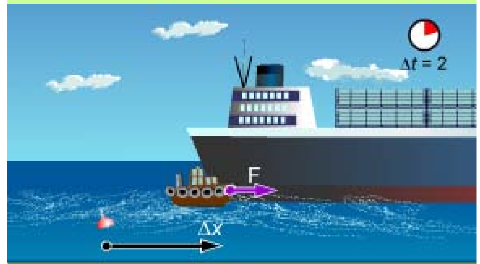
Equation 2 shows two other useful equations for power. Power can be expressed as the rate of change of energy, as you see in the first equation in Equation 2. Sometimes this is stated as "energy consumption", as in a 100-watt light bulb "consumes" 100 joules of energy each second. In other words, the light bulb converts 100 joules of electrical energy each second into other forms of energy, such as light and heat.

The companies that provide electrical power to homes measure each household's energy consumption in kilowatt·hours. You can check that this is a unit of energy. The companies multiply power (thousands of joules per second, or kilowatts) by time (hours). The result is that one kilowatt·hour equals 60 kilojoules, a unit of energy.

The second equation in Equation 2 shows that when there is a constant force **in the direction of an object's displacement**, the power can be measured as the product of the force and the velocity. This equation can be derived from the definition of work: $W = F\Delta x$. Dividing both sides of that equation by time yields power on the left (work divided by time). On the right side, dividing displacement by time yields velocity.

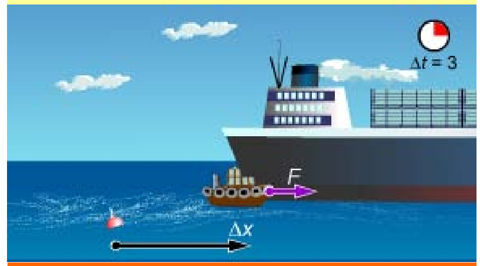
As with other rates of change, such as velocity or acceleration, we can consider average or instantaneous power. Average power is the total amount of work done over a period of time, divided by that time. Instantaneous power has the same definition, but the time interval must be a brief instant (more precisely, it is defined as the limit of the average power as the time interval approaches zero).

concept 1



Power
Rate of work

equation 1

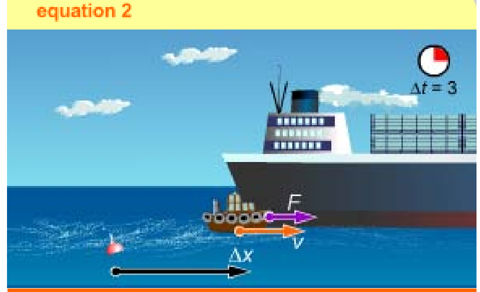


Power

$$\bar{P} = \frac{W}{\Delta t}$$

$\bar{P}$ = average power
W = work
Δt = time
Units: watts (W)

equation 2



Other power equations

$$\bar{P} = \frac{\Delta E}{\Delta t}$$

$$\bar{P} = F\,\bar{v}$$

E = energy
F = force
$\bar{v}$ = average velocity

example 1

Applying a force of 2.0×10^5 N, the tugboat moves the log boom 1.0 kilometer in 15 minutes. What is the tugboat's average power?

$W = F\Delta x$

$W = F\Delta x = (2.0\times10^5\text{ N})(1.0\times10^3\text{ m})$

$W = 2.0\times10^8\text{ J}$

$$\bar{P} = \frac{W}{\Delta t} = \frac{2.0\times10^8\text{ J}}{9.0\times10^2\text{ s}}$$

$\bar{P} = 2.2\times10^5\text{ W}$

7.13 - Potential energy

Potential energy: Energy related to the positions of and forces between the objects that make up a system.

Although the paint bucket in Concept 1 is not moving, it makes up part of a system that has a form of energy called potential energy. In general, potential energy is the energy due to the configuration of objects that exert forces on one other.

In this section, we focus on one form of potential energy, *gravitational potential energy*. The paint bucket and Earth make up a system that has this form of potential energy.

A *system* is some "chunk" of the universe that you wish to study, such as the bucket and the Earth. You can imagine a boundary like a bubble surrounding the system, separating it from the rest of the universe. The particles within a system can interact with one another via internal forces or fields. Particles outside the system can interact with the system via external forces or fields.

Gravitational potential energy is due to the gravitational force between the bucket and Earth. As the bucket is raised or lowered, its **change** in potential energy (ΔPE) equals the magnitude of its weight, mg, times its vertical displacement, Δh. (We follow the common convention of using Δh for change in height, instead of Δy.) The weight is the amount of force exerted on the bucket by the Earth (and vice versa). This formula is shown in Equation 1.

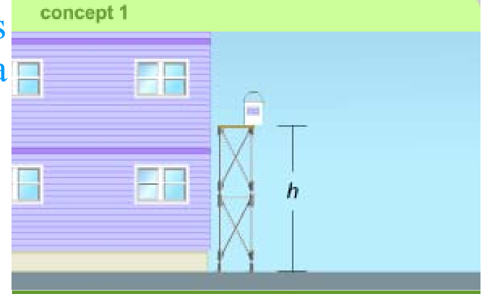
A change in PE can be positive or negative. The magnitude of weight is a positive value, but change in height can be positive (when the bucket moves up) or negative (when it moves down).

To define a system's PE, we must define a configuration at which the system has zero PE. Unlike kinetic energy, where zero KE has a natural value (when an object's speed is zero), the configuration with zero PE is defined by you, the physicist.

In the diagrams to the right, it is convenient to say the system has zero PE when the bucket is on the Earth's surface. This convention means its PE equals its weight times its height above the ground, mgh. Only the bucket's distance above the Earth, h, matters here; if the bucket moves left or right, its PE does not change.

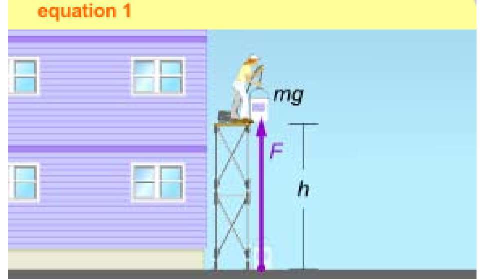
In Example 1, we calculate the paint bucket's gravitational potential energy as it sits on the scaffolding, four meters above the ground.

concept 1



Potential energy
Energy of position or configuration

equation 1



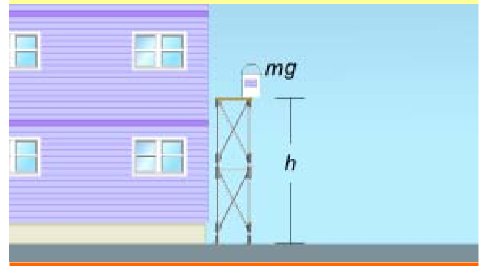
Change in gravitational potential energy

$\Delta PE = mg\Delta h$

PE = potential energy
mg = object's weight
Δh = vertical displacement

There are other types of potential energy. One you will frequently encounter is *elastic potential energy*, which is the energy stored in a compressed or stretched object such as a spring. As you may recall, this form of energy was present in the bow that was used to fire an arrow.

equation 2

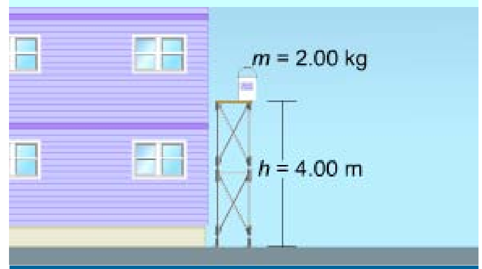


Gravitational potential energy

$PE = mgh$

$PE = 0$ when $h = 0$

example 1



What is the bucket's gravitational potential energy?

$PE = mgh$

$PE = (2.00 \text{ kg})(9.80 \text{ m/s}^2)(4.00 \text{ m})$

$PE = 78.4 \text{ J}$

7.14 - Work and gravitational potential energy

Potential energy is the energy of a system due to forces between the particles or objects that make up the system. It can be related to the amount of work done on a system by an external force. We will use gravitational force and gravitational potential energy as an example of this general principle. Our discussion applies to what are called conservative forces, a type of force we will later discuss in more detail.

The system we consider consists of two objects, the bucket and the Earth, illustrated to the right. The painter applies an external force to this system (via a rope) when she raises or lowers the bucket. The bucket starts at rest on the ground, and she raises it up and places it on the scaffolding. That means the work she does as she moves the bucket from its initial to its final position changes only its gravitational PE. The system's kinetic energy is zero at the beginning and the end of this process.

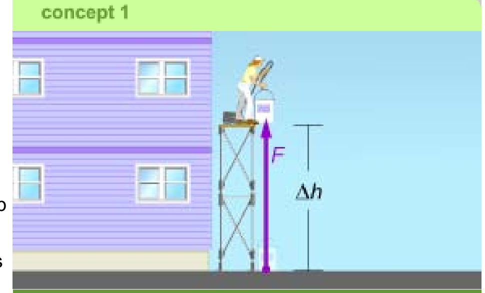
As she raises the bucket, the painter does work on it. She pulls the bucket up against the force of gravity, which is equal in magnitude to the bucket's weight, mg. She pulls in the direction of the bucket's displacement, Δh. The work equals the force multiplied by the displacement: $mg\Delta h$. The paint bucket's change in gravitational potential energy also equals $mg\Delta h$. The analysis lets us reach an important conclusion: The work done on the system, against gravity, equals the system's increase in gravitational potential energy.

Earlier, we stated the work-kinetic energy theorem: The net work done on a particle equals its change in kinetic energy. Here, where there is no change in kinetic energy, we state that the work done on a system equals its change in potential energy.

Are we confused? No. Work performed on a system can change its mechanical energy, which consists of its kinetic energy and its potential energy. Either or both of these forms of energy can change when work is applied to the system.

As the painter does work against the force of gravity, the force of gravity itself is also doing work. The work done by gravity is the negative of the work done by the painter. This means the work done by gravity is also the negative of the change in potential

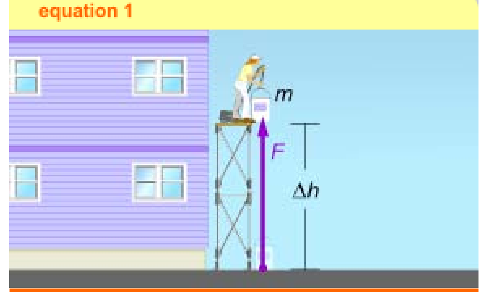
concept 1



Work and potential energy

Work equals change in energy

equation 1



Work done against gravity

energy, as seen in Equation 2.

Imagine that the painter drops the bucket from the scaffolding. Only the force of gravity does work on the bucket as it falls. The system has more potential energy when the bucket is at the top of the scaffolding than when it is at the bottom, so the work done by gravity has lowered the system's PE: the change in PE due to the work done by gravity is negative.

$$W = \Delta PE$$

W = work done against gravity
PE = potential energy of system

equation 2

Work done by gravity

$$W = -\Delta PE$$

W = work done by gravity
PE = potential energy of system

7.15 - Sample problem: potential energy and Niagara Falls

In its natural state, an average of 5.71×10^6 kg of water flowed per second over Niagara Falls, falling 51.0 m. If all the work done by gravity could be converted into electric power as the water fell to the bottom, how much power would the falls generate?

Variables

height of falls	$h = 51.0$ m
magnitude of acceleration due to gravity	$g = 9.80$ m/s^2
potential energy	PE
mass of water over falls per unit time	$m/t = 5.71\times10^6$ kg/s
power	P
work done by gravity	W

What is the strategy?

1. Use the definition of power as the rate of work done to define an equation for the power of the falls.
2. Use the fact that work done by gravity equals the negative of the change in gravitational potential energy to solve for the power.

Physics principles and equations

Power is the rate at which work is performed.

$$P = \frac{W}{\Delta t}$$

Change in gravitational PE

$$\Delta PE = mg\Delta h$$

Work done by gravity

$$W = -\Delta PE$$

Step-by-step solution

We start with the definition of power − work done per unit time − and then substitute in the definition of work done by gravity and the definition of gravitational potential energy to solve the problem.

Step	Reason
1. $P = \frac{W}{\Delta t}$	power equation
2. $P = \frac{-\Delta PE}{\Delta t}$	work done by gravity lowers PE
3. $P = \frac{-(mg\Delta h)}{\Delta t} = \frac{-mg(h_f - h_i)}{\Delta t}$ $P = \frac{mg(h_i - h_f)}{\Delta t}$	definition of gravitational potential energy
4. $P = \frac{m(9.80 \text{ m/s}^2)(51.0 \text{ m} - 0 \text{ m})}{\Delta t}$	enter values for g and h
5. $P = \left(5.71 \times 10^6 \ \frac{\text{kg}}{\text{s}}\right)(9.80 \text{ m/s}^2)(51.0 \text{ m})$	enter value for m/t
6. $P = 2.85 \times 10^9$ W	solve

This is the theoretical maximum power that could be generated. A real power plant cannot be 100% efficient.

7.16 - Interactive checkpoint: an elevator

An elevator with a mass of 515 kg is being pulled up a shaft at constant velocity. It takes the elevator 3.00 seconds to travel from floor two to floor three, a distance of 4.50 m. What is the average power of the elevator motor during this time? Neglect friction.

Answer:

$P =$ ______ W

7.17 - Work and energy

We have discussed work on a particle increasing its KE, and work on a system increasing its PE. Now we discuss what happens when work increases both forms of mechanical energy.

Because we are considering only KE and PE in this chapter, we can say the net work done on an object equals the change in the sum of its KE and PE. Positive work done on an object increases its energy; negative work decreases its energy.

Let's also consider what happens when an object does work, and how that affects the object's energy. Consider a soccer ball slamming into the hands of a goalie. The ball is doing work, forcing the goalie's hands backwards. The ball slows down; its energy decreases. Work done **by** an object **decreases** its energy. At the same time, this work on the goalie increases her energy. Work has transferred energy from one system (the ball) to another (the goalie).

We will use the scenario in Example 1 to show how both an object's KE and PE can change when work is done on it. A cannon shoots a 3.20

kg cannonball straight up. The barrel of the cannon is 2.00 m long, and it exerts an average force of 6,250 N while the cannonball is in the cannon. We will ignore air resistance. Can we determine the cannonball's velocity when it has traveled 125 meters upward?

As you may suspect, the answer is "yes".

The cannon does 12,500 J of work on the cannonball, the product of the force (6,250 N) and the displacement (2.00 m). (We assume the cannon does no work on the cannonball after it leaves the cannon.)

At a height of 125 meters, the cannonball's increase in PE equals $mg\Delta h$, or 3,920 J. Since a total of 12,500 J of work was done on the ball, the rest of the work must have gone into raising the cannonball's KE: The change in KE is 8,580 J.

Applying the definition of kinetic energy, we determine that its velocity at 125 m is 73.2 m/s. We could further analyze the cannonball's trip if we were so inclined. At the peak of its trip, all of its energy is potential since its velocity (and KE) there are zero. The PE at the top is 12,500 J. Again applying the formula $mg\Delta h$, we can determine that its peak height above the cannon is about 399 m.

concept 1

Work and energy
Work on system equals its change in total energy

example 1

The cannon supplies 6,250 N of force along its 2.00 m barrel. How much work does the cannon do on the cannonball?

$W = (F \cos \theta)\Delta x = F\Delta x$

$W = (6250\text{ N})(2.00\text{ m}) = 12{,}500\text{ J}$

example 2

What is the cannonball's velocity at 125 m? Its mass is 3.20 kg.

$W = \Delta PE + \Delta KE$

$W = mg\Delta h + \Delta KE$

$12{,}500\text{ J} = (3.20\text{ kg})(9.80\text{ m/s}^2)(125\text{ m}) + \Delta KE$

$\Delta KE = 8{,}580\text{ J}$

$\tfrac{1}{2} mv^2 = 8{,}580\text{ J}$

$v^2 = 2(8{,}580\text{ J})/(3.20\text{ kg})$

$v = 73.2\text{ m/s}$

7.18 - Conservative and non-conservative forces

Earlier, when discussing potential energy, we mentioned that we would explain conservative forces later. The concept of potential energy only applies to conservative forces.

Gravity is an example of a *conservative force*. It is conservative because the total work it does on an object that starts and finishes at the same point is zero. For example, if a 20 kg barbell is raised 2.0 meters, gravity does −40 J of work, and when the barbell is lowered 2.0 meters back to its initial position, gravity does +40 J of work. When the barbell is returned to its initial position, the sum of the work done by gravity on the barbell equals zero.

You can confirm this by considering the barbell's gravitational potential energy. Since that equals mgh, it is the same at the beginning and end because the height is the same. Since there is no change in gravitational PE, there is no work done by gravity on the barbell.

We illustrate this with the roller coaster shown in Concept 1. For now, we ignore other forces, such as friction, and consider gravity as the only force doing work on the roller coaster car. When the roller coaster car goes down a hill, gravity does positive work. When the roller coaster car goes up a hill, gravity does negative work. The sum of the work done by gravity on this journey equals zero.

When a roller coaster car makes such a trip, the roller coaster car travels on what is called a closed path, a trip that starts and stops at the same point. Given a slight push at the top of the hill, the roller coaster would make endless trips around the roller coaster track.

Kinetic friction and air resistance are two examples of *non-conservative forces*. These forces oppose motion, whatever its direction. Friction and air resistance do negative work on the roller coaster car, slowing it regardless of whether it is going uphill or downhill.

We show non-conservative forces at work in Concept 2. The roller coaster glides down the hill, but it does not return to its initial position because kinetic friction and air resistance dissipate some of its energy as it goes around the track. The presence of these forces dictates that net work must be done on the roller coaster car by some other force to return it to its initial position. A mechanism such as a motorized pulley system can accomplish this.

A way to differentiate between conservative and non-conservative forces is to ask: Does the amount of work done by the force depend on the path?

Consider only the force of gravity, a conservative force, as it acts on the skier shown in Concept 3. When considering the work done by gravity, it does not matter in terms of work and energy whether the skier goes down the longer, zigzag route (path A), or the straight route (path B). The work done by gravity is the same along either path. All that matters are the locations of the initial and final points of the path. The conservative force is *path independent*.

However, with non-conservative forces, the path does influence the amount of work done by a force. Consider the skier in the context of kinetic friction, a non-conservative force. The amount of force of kinetic friction is the same along either route, but it acts along a greater distance if the skier chooses the longer route. The amount of work done by the force of kinetic friction increases with the path length. A non-conservative force is *path dependent*.

concept 1

Conservative force
A force that does no work on closed path

concept 2

Non-conservative force
A force that does work on closed path

concept 3

Effect of path on work and energy
Conservative forces: work does not depend on path
Non-conservative forces: work depends on path

7.19 - Conservation of energy

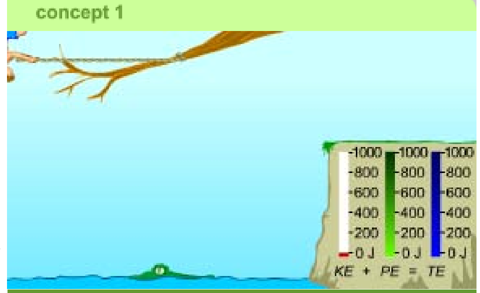
Conservation of energy: The total energy in an isolated system remains constant.

Energy never disappears. It only changes form and transfers between objects.

To illustrate this principle, we use the boy to the right who is swinging on a rope. We consider his mechanical energy, the sum of his kinetic and gravitational potential energy. When he jumps from the riverbank and swings toward the water, his gravitational potential energy becomes kinetic energy. The decrease in gravitational PE (shown in the gauge labeled PE in Concept 1) is matched by an increase in his kinetic energy (shown in the gauge labeled KE). Ignoring air resistance or any other non-conservative forces, the sum of KE and PE is a constant at any point. The total amount of energy (labeled TE, for total energy) stays the same. The total energy is conserved; its amount does not change.

The law of conservation of energy applies to an isolated system. An *isolated system* is

concept 1



Conservation of energy
Total energy in isolated system stays constant

one that has no interactions with its environment. The particles within the system may interact with one another, but no net external force or field acts on an isolated system.

Only external forces can change the total energy of a system. If a giant spring lifts a car, you can say the spring has increased the energy of the car. In this case, you are considering the spring as supplying an external force and not as part of the system. If you include the spring in the system, the increase in the energy of the car is matched by a decrease in the potential energy contained of the spring, and the total energy of the system remains the same. For the law of conservation of energy to apply, there can be no non-conservative forces like friction within the system.

The law of conservation of energy can be expressed mathematically, as shown in Equation 1. The equation states that an isolated system's total energy at any final point in time is the same as its total energy at an initial point in time. When considering mechanical energy, we can state that the sum of the kinetic and potential energies at some final moment equals the sum of the kinetic and potential energies at an initial moment.

In the case of the boy on the rope, if you know his mass and height on the riverbank, you can calculate his gravitational potential energy. In this example, rather than saying his PE equals zero on the ground, we say it equals zero at the bottom of the arc. This simplifies matters. Using the law of conservation of energy, you can then determine what his kinetic energy, and therefore his speed, will be when he reaches the bottom of the arc, nearest to the water, since at that point all his energy is kinetic.

Let's leave the boy swinging for a while and switch to another example: You drop a weight. When the weight hits the ground it will stop moving. At this point, the weight has neither kinetic energy nor potential energy because it has no motion and its height off the Earth's surface is zero. Does the law of conservation of energy still hold true?

Yes, it does, although we need to broaden the forms of energy included in the discussion. With careful observation you might note that the ground shakes as the weight hits it (more energy of motion). The weight and the ground heat up a bit (thermal energy). The list can continue: energy of the motion of flying dirt, the energy of sound and so on. The amount of mechanical energy does decline, but when you include all forms of energy, the overall energy stays constant.

There is a caveat to the law of conservation of energy. Albert Einstein demonstrated that there is a relationship between mass and energy. Mass can be converted into energy, as it is inside the Sun or a nuclear reactor, and energy can be converted into mass. It is the sum of mass and energy that remains constant. Our current focus is on much less extreme situations.

Using the principle of conservation of energy can have many practical benefits, as automotive engineers are now demonstrating. When it comes to energy and cars, the focus is often on how to cause the car to accelerate, how fast they will reach say a speed of 100 km/h.

Of course, cars also need to slow down, a task assigned to the brakes. As conventional cars brake, the energy is typically dissipated as heat as the brake pads rub on the rotors. Innovative new cars, called hybrids, now capture some of the kinetic energy and convert it to chemical energy stored in batteries or mechanical energy stored in flywheels. The engine then recycles that energy back into kinetic energy when the car needs to accelerate, saving gasoline.

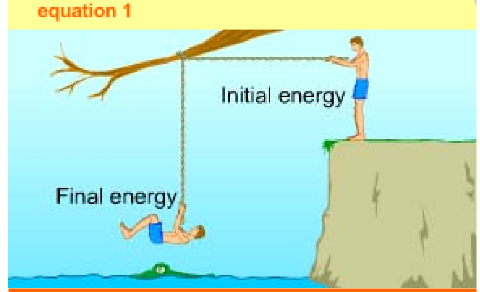
concept 2

Conservation of energy
PE transforms to *KE* ...

concept 3

Conservation of energy
... but total energy remains the same

equation 1



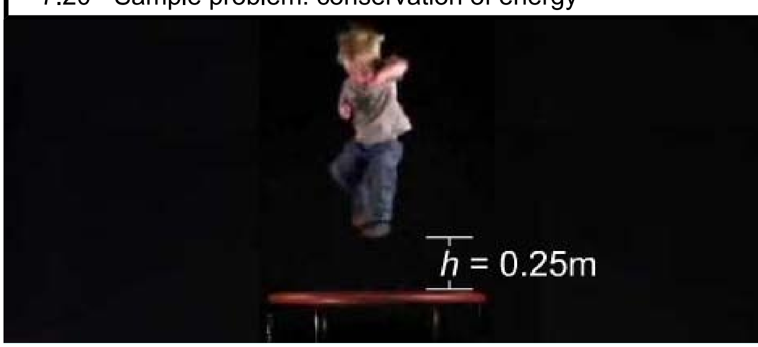
Conservation of energy

$$E_f = E_i$$

$$PE_f + KE_f = PE_i + KE_i$$

E = total energy
KE = kinetic energy
PE = potential energy

7.20 - Sample problem: conservation of energy



Sam is at the peak of his jump. Calculate Sam's speed when he reaches the trampoline's surface.

Sam is jumping up and down on a trampoline. He bounces to a maximum height of 0.25 m above the surface of the trampoline. How fast will he be traveling when he hits the trampoline? We define Sam's potential energy at the surface of the trampoline to be zero.

Variables

Sam's height at peak	$h = 0.25$ m
Sam's speed at peak	$v_{peak} = 0$ m/s
Sam's speed at bottom	v

What is the strategy?

1. Use the law of conservation energy, to state that Sam's total energy at the peak of his jump is the same as his total energy at the surface of the trampoline. Simplify this equation, using the facts that his kinetic energy is zero at the peak, and his potential energy is zero at the surface of the trampoline.
2. Solve the resulting equation for his speed at the bottom.

Physics principles and equations

The definition of gravitational potential energy

$$PE = mgh$$

The definition of kinetic energy

$$KE = ½\, mv^2$$

Total energy is conserved in this isolated system.

$$E_f = E_i$$

Step-by-step solution

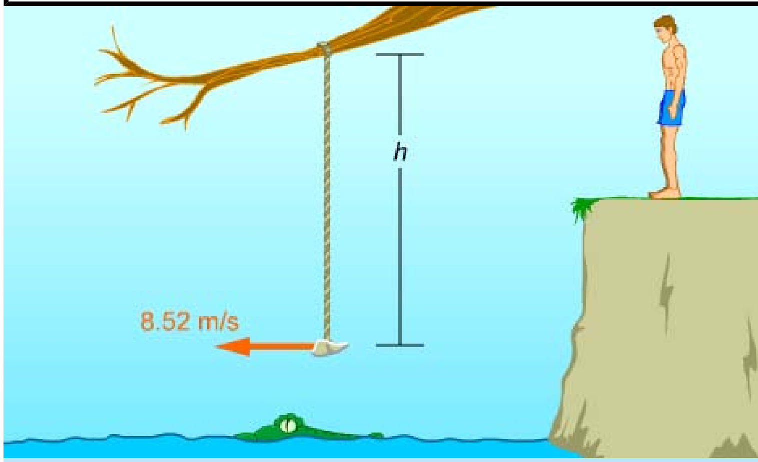
We start by stating the law of conservation of energy in equation form, and then adapting it to fit the specifics of Sam's trampoline jump.

Step	Reason
1. $E_f = E_i$	law of conservation of energy
2. $KE_f + PE_f = KE_i + PE_i$	energy is mechanical energy
3. $KE_f + 0 = 0 + PE_i$	enter values
4. $KE_f = PE_i$	simplify

Now we have a simpler equation for Sam's energy. The task now is to solve it for the one unknown variable, his speed at the surface of the trampoline.

Step	Reason
5. $KE_f = PE_i$	state equation again
6. $½\, mv^2 = mgh$	definitions of KE, PE
7. $v = \sqrt{2gh}$	solve for v
8. $v = \sqrt{2\,(9.80\ \text{m/s}^2)(0.25\ \text{m})}$	enter values
9. $v = 2.2$ m/s	evaluate

7.21 - Interactive checkpoint: conservation of energy



A boy releases a pork chop on a rope. The chop is moving at a speed of 8.52 m/s at the bottom of its swing. How much higher than this point is the point from which the pork chop is released? Assume that it has no initial speed when starting its swing.

Answer:

h = ________ m

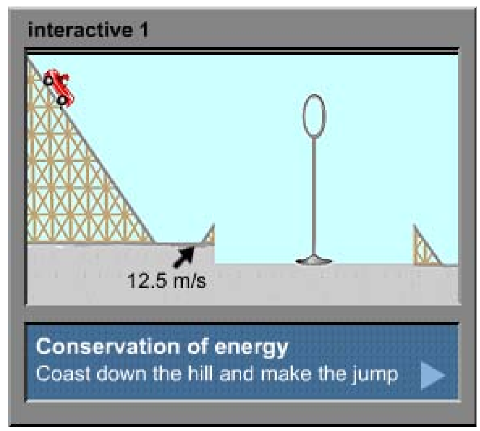
7.22 - Interactive problem: conservation of energy

The law of conservation of energy states that the total energy in an isolated system remains constant. In the simulation on the right, you can use this law and your knowledge of potential and kinetic energies to help a soapbox derby car make a jump.

A soapbox derby car has no engine. It gains speed as it rolls down a hill. You can drag the car to any point on the hill. A gauge will display the car's height above the ground. Release the mouse button and the car will fly down the hill.

In this interactive, if the car is traveling 12.5 m/s at the bottom of the ramp, it will successfully make the jump through the hoop. Too slow and it will fall short; too fast and it will overshoot.

You can use the law of conservation of energy to figure out the vertical position needed for the car to nail the jump.

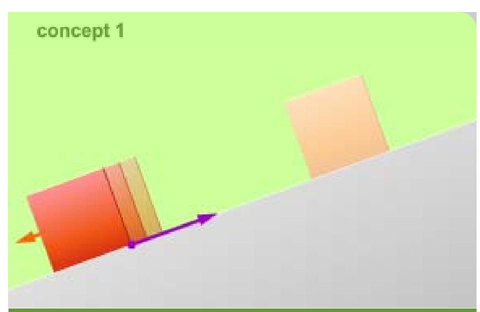


7.23 - Friction and conservation of energy

In this section, we show how two principles we have discussed can be combined to solve a typical problem. We will use the principle of conservation of energy and how work done by an external force affects the total energy of a system to determine the effect of friction on a block sliding down a plane.

Suppose the 1.00 kg block shown to the right slides down an inclined wooden plane. Since the block is released from rest, it has no initial velocity. It loses 2.00 meters in height as it slides, and it slides 6.00 meters along the surface of the inclined plane. The force of kinetic friction is 2.00 N. You want to know the block's speed when it reaches the bottom position.

To solve this problem, we start by applying the principle of conservation of energy. The block's initial energy is all potential, equal to the product of its mass, g and its height (mgh). At a height of 2.00 meters, the block's PE equals 19.6 J. The potential energy will be zero when the block reaches the bottom of the plane. Ignoring friction, the PE of the block at the top equals its KE at the bottom.



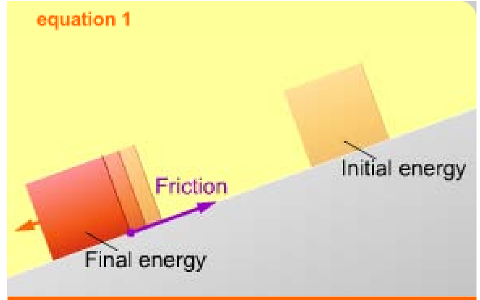
Effect of non-conservative forces
Reduce object's energy

Now we will factor in friction. The force of friction opposes the block's motion down the inclined plane. The work it does is negative, and that work reduces the energy of the block. We calculate the work done by friction on the block as the force of friction times the displacement along the plane, which equals −12.0 J. The block's energy at the top (19.6 J) plus the −12.0 J means the block has 7.6 J of kinetic energy at the bottom. Using the definition of kinetic energy, we can conclude that the 1.00 kg block is moving at 3.90 m/s.

You can also calculate the effect of friction by determining how fast the block would be traveling if there were no friction. All 19.6 J of PE would convert to KE, yielding a speed of 6.26 m/s. Friction reduces the speed of the block by approximately 38%.

In general, non-conservative forces like friction and air resistance are *dissipative forces*: They reduce the energy of a system. They do negative work since they act opposite the direction of motion. (There are a few cases where they do positive work, such as when the force of friction causes something to move, as when you step on a moving sidewalk.)

equation 1

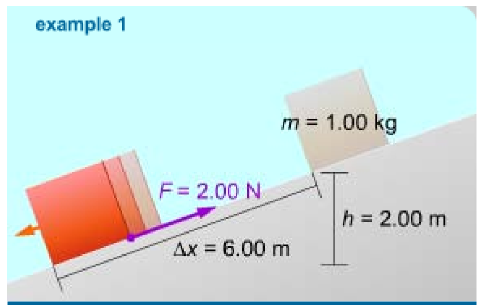


Effect of non-conservative forces

$$W_{nc} = E_f - E_i$$

W_{nc} = work by non-conservative force
E_f = final energy
E_i = initial energy

example 1



What is the block's kinetic energy at the bottom?

$W_{nc} = E_f - E_i$

$(F \cos \theta)\Delta x = KE_f - PE_i$

$KE_f = PE_i + (F \cos \theta \, \Delta x)$

$KE_f = (mgh) + (F \cos \theta \, \Delta x)$

$KE_f = (1.00 \text{ kg} \times 9.80 \text{ m/s}^2 \times 2.00 \text{ m}) + (2.00 \text{ N} \times \cos 180° \times 6.00 \text{ m})$

$KE_f = 19.6 \text{ J} + (-12.0 \text{ J})$

$KE_f = 7.60 \text{ J}$

7.24 - Interactive problem: a non-conservative force

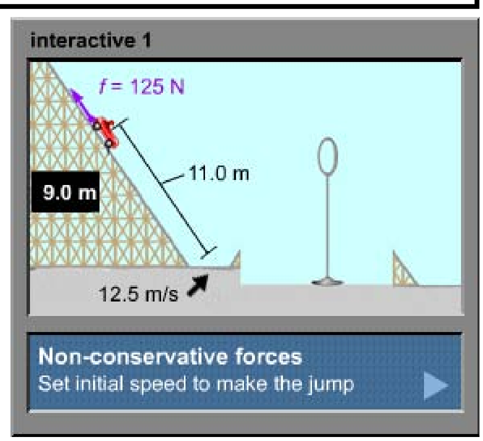
You may have helped the car make this jump in a previous simulation.

Here is the same scenario, only now we have added the non-conservative force of kinetic friction to the ramp. A constant 125 N frictional force opposes the soapbox car's motion as it goes down the hill. In order to make the jump, you will have to take the force of friction into account.

A 100.0 kg car is situated at a vertical height of 9.0 meters above the bottom of the hill. Your car needs to be traveling at 12.5 m/s at the bottom of the ramp to fly through the hoop. Friction will oppose the car's motion as it rolls down the ramp.

To correct for the energy lost to friction, you must supply an initial speed to the car so that it makes the jump. Calculate that value, enter it in the simulation to the nearest tenth of a meter per second, and press GO to test your answer.

interactive 1



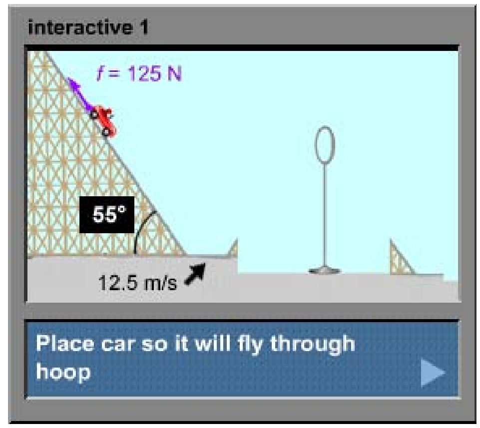
7.25 - Interactive problem: a non-conservative force, part II

Here is another version of the interactive problem in which you place the car on the hill at the correct height to make the jump.

Again, a constant 125 N frictional force opposes the soapbox car's motion as it goes down the hill. In order to make the jump, you will have to take this force into account. Once again, the mass of the car is 100.0 kg. As shown, the angle of the hill is $55°$, measured from the negative *x* axis.

You can set the car at any point on the hill. A gauge displays the car's vertical height above the ground. Its initial speed will be zero. If you are traveling at 12.5 m/s at the bottom, you will successfully make the jump. If you are too slow you will fall short and if you are too fast you will fly too high.

HINT: To solve the problem, you must determine a way to express the distance the car travels in terms of its vertical height.



7.26 - Review of forces, work and energy

In this chapter we have discussed the work done when a force is exerted on a particle (the work-kinetic energy theorem). We have discussed work and energy with respect to a system of objects (potential energy). We have also covered conservative and non-conservative forces.

We further categorized forces by stating that some are *external forces*, forces from a source outside the objects that make up the system. For example, we talked about a foot applying force to a soccer ball, and a painter hoisting up a paint bucket. In both these examples, the foot and painter are considered external to the system, which consists either of a single particle (the ball) or multiple objects (the bucket and the Earth).

In contrast, other forces are *internal forces* in a system. In the bucket/Earth system, for example, the force of gravity is an internal force. It arises from the objects that make up the system.

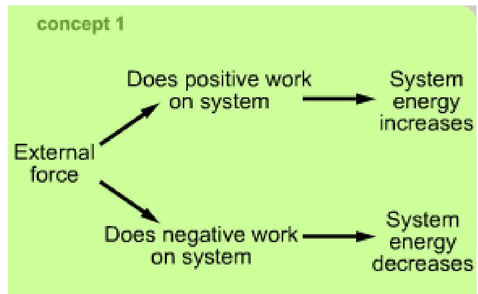
In this section, we review and summarize the effect on mechanical energy from all these types of forces: external and internal, conservative and non-conservative. We want to consider how the work done by these various types of forces affects the mechanical energy of a system.

We will start with external forces, and consider the effect of the work done by an external force on the total energy of a system. Any net external force acting on a particle or system changes the system's energy.

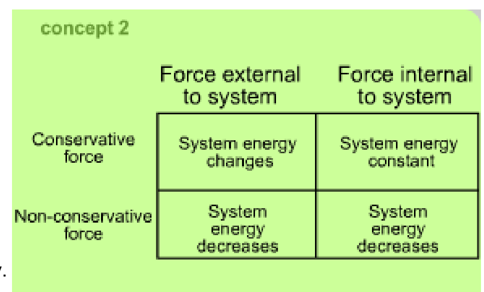
Positive work done on a system by an external force increases the system's total energy, and negative work done on a system by an external force decreases its total energy. This is illustrated in the diagram in Concept 1.

If the system consists of one particle, then the work equals the change in kinetic energy. This is the work-kinetic energy theorem. A single particle cannot have potential energy, so positive work on the particle increases its KE, and negative work done on the particle (or work done by the particle) decreases its KE.

The work done on any system by an external force changes the system's total mechanical energy. Let's consider a system that consists of an apple and the Earth. Positive work may increase the PE (you lift the apple upwards at a constant rate), or KE (you run faster and faster with the apple held at a constant height), or both (you throw an initially stationary apple skyward).



Summary of effect of external forces on system energy



Summary of effect of conservative and non-conservative forces on system energy

You need to be careful of the sign of the work done by considering whether the force is in the direction of the displacement (positive work) or the opposite direction (negative work). If you throw a ball, you increase its energy, and when you catch it, you decrease its energy.

Non-conservative forces decrease the mechanical energy of a system. (There are scenarios where they can be considered as increasing the mechanical energy, but we will ignore them here.) If you slide a block down a plane, the non-conservative forces of kinetic friction and air resistance act in the opposite direction of the block's displacement. This means they do negative work, and reduce the mechanical energy of the system.

Now let's consider the effect of internal forces on the energy of a system. We will start with internal conservative forces. These forces do not change the total mechanical energy of a system. Consider a system consisting of a block, an inclined plane, and the Earth. The block is sliding down the plane. The force of gravity is conservative, and the decrease in gravitational potential energy as the block slides down the plane is matched by an increase in kinetic energy. This is the law of conservation of energy. The conservative force of gravity does not change the total mechanical energy of the system.

We will foreshadow thermodynamics here. The force of friction will increase the temperature of the block and plane. It increases the *internal* *energy* of the system, the random motion of the particles that make up the block and plane. Several famous experiments showed the relationship between mechanical energy and internal energy by measuring changes in temperature. The total energy, which is the sum of the internal energy, PE and $KE,$ remains constant. But in mechanics, we only track the mechanical energy of the system.

We can also discuss the relationship of the work done by an internal conservative force and the system's potential energy. The work done by an internal force equals the negative of the change in the system's corresponding PE. The work done by the force of gravity on the block equals the negative of the change in gravitational PE of the block/plane/Earth system. The reduction in PE is matched by an increase in KE.

We summarize the roles of conservative and non-conservative forces, and internal and external forces, in the table in Concept 2.

Although we present this summary and these tables to summarize important points, we also encourage you to apply your physics intuition. If you do work on an object, causing it to move faster, you have increased its energy. If friction slows something down, it has reduced the object's energy. Keeping in mind the realities of the situation that you are dealing with is the best tool of all.

7.27 - Interactive summary problem: work, energy and power

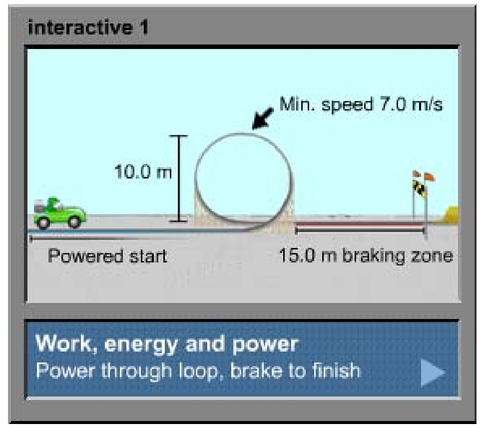
In this simulation, you are driving a motorized go-kart. There is no force of friction in this simulation. You need to set the speed of the go-kart so it goes fast enough to make it all the way around the loop-the-loop, then brake it hard enough that it stops before hitting the hay bales just past the finish line.

You control the power of the go-kart as it first accelerates, and then set the amount of force required to brake the car.

To solve this problem, you need to make use of the major concepts of this chapter: work, power and energy.

Set the power in watts to the nearest 100 W for your 98 kg go-kart so that if the power is applied for 2.0 seconds, the cart will have enough energy to make it around the loop-the-loop. The minimum speed at the top of the loop is 7.0 m/s. The loop is 10.0 meters high.

As it exits the loop, your car is 15.0 meters from the finish line. Just beyond the finish line are hay bales that, if you hit them, make for a very scratchy ending. Calculate the frictional braking force to the nearest 10 N that will allow the cart to make it to the finish line but stop before hitting the bales. In other words, bring the cart to a stop in exactly 15.0 meters.



7.28 - Gotchas

You are asked to push two wheelbarrows up a hill. One wheelbarrow is empty, and you are able to push it up the hill in one minute. The other is filled with huge rocks, and even after you push it for an hour, you cannot budge it. In which case do you do more work on the wheelbarrow? You do more work on the empty wheelbarrow because it is the only one that moves. You do no work on the wheelbarrow filled with rocks because you do not move it; its displacement is zero.

Two 1/4 kg cheeseburgers are moving in opposite directions. One is rising at three m/s, the other is falling at three m/s. Which has more kinetic energy? They are the same. The direction of velocity does not matter for kinetic energy; only the magnitude of velocity (the speed) matters.

You start on a beach. You go to the moon. You come back. You go to Hollywood. You then go to the summit of Mount Everest. Have you increased your gravitational potential energy more than if you had climbed to this summit without all the other side trips? No, the increase in energy is the same. In both cases, the increase in gravitational potential energy equals your mass times the height of Mount Everest times g.

You are holding a cup of coffee at your desk, a half-meter above the floor. You extend your arm laterally out a nearby third-story window so that the cup is suspended 10 meters above the ground. Have you increased the cup's gravitational potential energy? No. Assume you choose the floor as the zero potential energy point. The potential energy is the same because the cup remains the same distance above the floor.

An apple falls to the ground. Because the Earth's gravity did work on it, the apple's energy has increased. It depends on how you define "the system." If you said the apple had gravitational potential energy, then you are including the Earth as part of the system. In this case, the Earth's gravity is not an external force. Energy is conserved: The system's decreased PE is matched by an increase in KE. You could also say "the system" is solely the apple. In that case, it has no PE, since PE requires the presence of a force between objects in a system, and this system has only one object. Gravity is now an external force acting on the apple, and it does increase the apple's KE. This is not, perhaps, the typical way to think about it, but it is valid.

It is impossible for a system to have negative PE. Wrong: Systems can have negative PE. For instance, if you define the system to have zero PE when an object is at the Earth's surface, the object has negative PE when it is below the surface. Its PE has decreased from zero, so it must be negative. Negative PE is common in some topics, such as orbital motion.

7.29 - Summary

Work is the product of the force on an object and its displacement in the direction of that force. It is a scalar quantity with units of joules (1 J = 1 kg·m^2/s^2).

Work and several other scalar quantities can be computed by taking the dot product of two vectors. The dot product is a scalar equal to the product of the magnitudes of the two vectors and the cosine of the angle between them. Loosely, it tells you how much of one vector is in the direction of another.

Energy is a property of an object or a system. It has units of joules and is a scalar quantity. Energy can transfer between objects and change forms. Work on an object or system will change its energy.

One form of energy is kinetic energy. It is the energy possessed by objects in motion and is proportional to the object's mass and the square of its speed.

The work-kinetic energy theorem states that the work done on a particle or an object modeled as a particle is equal to its change in kinetic energy. Positive work increases the energy, while negative work decreases it.

Power is work divided by time. The unit of power is the watt (1 W = 1 J/s), a scalar quantity. It is often expressed as a rate of energy consumption or output. For example, a 100-watt light bulb converts 100 joules of electrical energy per second into light and heat.

Another form of energy is potential energy. It is the energy related to the positions of the objects in a system and the forces between them. Gravitational potential energy is an object's potential energy due to its position relative to a body such as the Earth.

Forces can be classified as conservative or non-conservative. An object acted upon only by conservative forces, such as gravitational and spring forces, requires no *net* work to return to its original position. An object acted upon by non-conservative forces, such as kinetic friction, will not return to its initial position without additional work being done on it.

When only conservative forces are present, the work to move an object between two points does not depend on the path taken. The work is path independent. When non-conservative forces are acting, the work does depend on the path taken, and the work is path dependent.

When work is being done by a conservative force within a system, the force can be calculated as the negative of the derivative of the potential energy curve with respect to displacement.

The law of conservation of energy states that the total energy in an isolated system remains constant, though energy may change form or be transferred from object to object within the system.

Mechanical energy is conserved only when there are no non-conservative forces acting in the system. When a non-conservative force such as friction is present, the mechanical energy of the system decreases. The law of conservation of energy still holds, but we have not yet learned to account for the other forms into which the mechanical energy might be transformed, such as thermal (heat) energy.

Equations

$$\mathbf{A} \cdot \mathbf{B} = AB \cos \theta$$

$$W = \mathbf{F} \cdot \Delta\mathbf{x} = (F \cos \theta)\Delta x$$

$$KE = \tfrac{1}{2} mv^2$$

$$W = \Delta KE, \text{ for a particle}$$

$$P = \frac{W}{\Delta t} = \frac{\Delta E}{\Delta t}$$

$$\Delta PE = mg\Delta h$$

$$E_{\mathrm{f}} = E_{\mathrm{i}}$$

Chapter 7 Problems

Conceptual Problems

C.1 You take your backpack out of the car, hike to the top of a nearby mountain, and return with your backpack to your car at the bottom. Have you changed the amount of your backpack's mechanical energy?

○ Yes ○ No

C.2 There are two slides at the park between which you are deciding. Both start at a height of 6 m. One is short and steep while the other is long and shallow. If both slides are frictionless, which one should you choose if you want to be moving as fast as possible at the bottom?

i. Short and steep
ii. Long and shallow
iii. It doesn't matter

C.3 A skydiver jumps out of a plane. Does the mechanical energy of the jumper change while she is falling? Ignore the effects of air resistance.

○ Yes ○ No

C.4 At the gym, your trainer places a 8.0 kg dumbbell in each of your outstreched hands. You hold them there for 20 seconds. Have you done any work during those 20 seconds?

○ Yes ○ No

C.5 Two boxes of the same mass are lifted to the same height. Does it necessarily take the same amount of power to lift each box?

○ Yes ○ No

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) If you release the car at a higher point, will its speed at the bottom of the hill be greater, the same, or less? (b) How high will you have to drag the car to have it just reach the summit of the other hill: to the same height, higher or lower?

(a) i. Greater
ii. The same
iii. Less

(b) i. Higher than the other hill
ii. To the same height as the other hill
iii. Lower than the other hill

Section 1 - Work

1.1 An airline pilot pulls her 12.0 kg rollaboard suitcase along the ground with a force of 25.0 N for 10.0 meters. The handle she pulls on makes an angle of 36.5 degrees with the horizontal. How much work does she do over the ten-meter distance?

__________ J

1.2 A parent pushes a baby stroller from home to daycare along a level road with a force of 34 N directed at an angle of 30° below the horizontal. If daycare is 0.83 km from home, how much work is done by the parent?

__________ J

1.3 A horizontal net force of 75.5 N is exerted on a 47.2 kg sofa, causing it to slide 2.40 meters along the ground. How much work does the force do?

__________ J

1.4 Charlie pulls horizontally to the right on a wagon with a force of 37.2 N. Sara pulls horizontally to the left with a force of 22.4 N. How much work is done on the wagon after it has moved 2.50 meters to the right?

__________ J

Section 2 - Dot product

2.1 Vector **v** has a magnitude of 44 m and vector **u** has a magnitude of 77 m. The angle between **v** and **u** is 150°. What is **v**·**u**?

__________ m^2

2.2 Find the dot product of (2, 3) and (4, −7).

____

2.3 The magnitude of **u** is 5.0, the magnitude of **v** is 7.0 and **u·v** is 28. What is the angle between **u** and **v**? Choose the positive solution between 0 and 180 degrees.

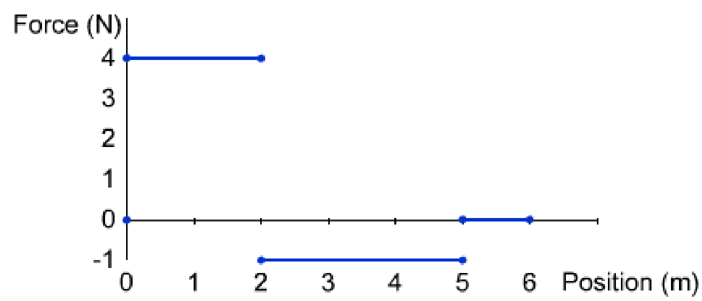
__________ °

2.4 The dot product obeys the distributive law: (**A** + **B**) · **C** = **A·C** + **B·C**. Show this is true for **A** = (5, 4), **B** = (1, −2), **C** = (−3, −3).

Section 3 - Work done by a variable force

3.1 The graph shown describes a certain force that is exerted on an object, as a function of the position of the object. How much work is done by this force as the object moves from the position 0.0 m to 6.0 m?

__________ J



Section 6 - Kinetic energy

6.1 You are about shoot two identical cannonballs straight up into the air. The first cannonball has 7.0 times as much initial velocity as the second. How many times higher will the first cannonball go compared to the second?

__________ times higher

6.2 What is the change in kinetic energy of a baseball as it accelerates from rest to 45.0 m/s? The mass of a baseball is 145 grams.

__________ J

6.3 A bullet of mass 10.8 g leaves a gun barrel with a velocity of 511 m/s. What is the bullet's kinetic energy?

__________ J

6.4 A 0.50 kg cream pie strikes a circus clown in the face at a speed of 5.00 m/s and stops. What is the change in kinetic energy of the pie?

__________ J

Section 7 - Work-kinetic energy theorem

7.1 A net force of 1.6×10^{-15} N acts on an electron over a displacement of 5.0 cm, in the same direction as the net force. (a) What is the change in kinetic energy of the electron? (b) If the electron was initially at rest, what is the speed of the electron? An electron has a mass of 9.1×10^{-31} kg.

(a) __________ J
(b) __________ m/s

7.2 A proton is moving at 425 m/s. (a) How much work must be done on it to stop it? (A proton has a mass of 1.67×10^{-27} kg.) (b) Assume the net braking force acting on it has magnitude 8.01×10^{-16} N and is directed opposite to its initial velocity. Over what distance must the force be applied? Watch your negative signs in this problem.

(a) __________ J
(b) __________ m

7.3 A hockey stick applies a constant force over a distance of 0.121 m to an initially stationary puck, of mass 152 g. The puck moves with a speed of 51.0 m/s. With what force did the hockey stick strike the puck?

__________ N

7.4 At the Aircraft Landing Dynamics Facility located at NASA's Langley Research Center in Virginia, a water-jet nozzle propels a 46,000 kg sled from zero to 400 km/hour in 2.5 seconds. (This sled is equipped with tires that are being tested for the space shuttle program, which are then slammed into the ground to see how they hold up.) (a) Assuming a constant acceleration for the sled, what distance does it travel during the speeding-up phase? (b) What is the net work done on the sled during the time interval?

(a) __________ m
(b) __________ J

7.5 A 25.0 kg projectile is fired by accelerating it with an electromagnetic rail gun on the Earth's surface. The rail makes a 30.0 degree angle with the horizontal, and the gun applies a 1250 N force on the projectile for a distance of 7.50 m along the rail. (a) Ignoring air resistance and friction, what is the net work done on the projectile, by all the forces acting on it, as it moves 7.50 m along the rail? (b) Assuming it started at rest, what is its speed after it has moved the 7.50 m?

(a) __________ J
(b) __________ m/s

7.6 The force of gravity acts on a 1250 kg probe in outer space. It accelerates the probe from a speed of 225 m/s to 227 m/s over a distance of 6750 m. How much work does gravity do on the probe?

__________ J

Section 10 - Interactive problem: work-kinetic energy theorem

10.1 Use the information given in the interactive problem in this section to answer the following question. What force is required for the skier to make the jump? Test your answer using the simulation.

__________ N

Section 12 - Power

12.1 How much power would be required to hoist a 48 kg couch up to a 22 m high balcony in 5.0 seconds? Assume it starts and ends at rest.

__________ W

12.2 Stuntman's Freefall, a ride at Six Flags Great Adventure in New Jersey, stands 39.6 meters high. Ignoring the force of friction, what is the minimum power rating of the motor that raises the 1.20×10^5 kg ride from the ground to the top in 10.0 seconds at a constant velocity?

__________ W

12.3 Lori, who loves to ski, has rigged up a rope tow to pull herself up a local hill that is inclined at an angle of 30.0 degrees from the horizontal. The motor works against a retarding frictional force of 100 N. If Lori has a mass of 60.5 kg, and the power of the motor is 1350 W, at what speed can the motor pull her up the hill?

__________ m/s

12.4 A power plant supplies 1,100 megawatts of power to the electric grid. How many joules of energy does it supply each second?

__________ J

12.5 How much work does a 5.00 horsepower outboard motor do in one minute? State your answer in joules.

__________ J

12.6 The Porsche® 911 GT3 has a 380 hp engine and a mass of 1.4×10^3 kg. The car can accelerate from 0 to 100 km/h in 4.3 seconds. What percentage of the power supplied by the engine goes into making the car move? Assume that the car's acceleration is constant and that there are 746 Watts/hp.

__________ %

Section 13 - Potential energy

13.1 You get paid $1.25 per joule of work. How much do you charge for moving a 10.0 kg box up to a shelf that is 1.50 meters off the ground?

$ __________

Section 14 - Work and gravitational potential energy

14.1 You lower a 2.50 kg textbook (remember when textbooks used to be made out of paper instead of being digital?) from a height of 1.85 m to 1.50 m. What is its change in potential energy?

__________ J

14.2 The reservoir behind the Grand Coulee dam in the state of Washington holds water with a total gravitational potential energy of 1×10^{16} J, where the reference point for zero potential energy is taken as the height of the base of the dam. (a) Suppose that the dam released all its water, which flowed to form a still pool at the base of the dam. What would be the change in the gravitational potential energy of the Earth-dam-water system? (b) What work was done by the gravitational force? (Part of this work is ordinarily used to turn electric generators.) (c) How much work would it take to pump all the water back up into the reservoir?

(a) __________ J
(b) __________ J
(c) __________ J

14.3 What is the change in the gravitational potential energy of a Boeing 767 jet as it soars from the runway up to a cruising altitude of 10.2 km? Assume its mass is a constant 2.04×10^5 kg.

__________ J

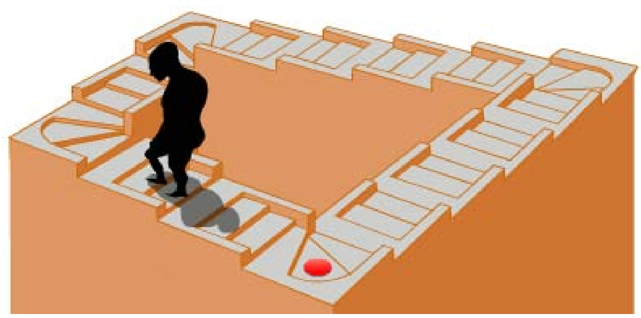
14.4 A weightlifter raises a 115 kg weight from the ground to a height of 1.95 m in 1.25 seconds. What is the average power of this maneuver?

__________ W

14.5 A small domestic elevator has a mass of 454 kg and can ascend at a rate of 0.180 m/s. What is the average power that must be supplied for the elevator to move at this rate?

__________ W

14.6 You see an optical illusion of an ever-upward spiral staircase. The climber trudges up and up and never gets anywhere, going in circles instead. Suppose the staircase is provided with a narrow ramp, allowing the tired stair-climber to push a wheelbarrow up the stairs. The loaded wheelbarrow weighs 300.0 N, and the ramp makes an angle of 15.0° with the horizontal, all along its length. The ramp consists of four straight sections, with slant lengths 12.0 m, 8.0 m, 20.0 m, and 20.0 m.



How much work does the climber do on the wheelbarrow when he pushes it up the ramp from the red marker, all the way around the loop, and (supposedly) back to the red marker again? An ordinary inclined-plane computation will give an accurate value for the work. (In the illusory illustration, the fact that he ends up where he started means that, impossibly, he does NO work.)

__________ N · m

Section 17 - Work and energy

17.1 A firm bills at the rate of \$1.00 per 125 J of work. It bills you \$45.00 for carrying a sofa up some stairs. Their workers moved the sofa up 3 flights of stairs, with each flight being 4.50 meters high. What is the mass of the sofa?

__________ kg

17.2 An engine supplies an upward force of 9.00 N to an initially stationary toy rocket, of mass 54.0 g, for a distance of 25.0 m. The rocket rises to a height of 339 meters before falling back to the ground. What was the magnitude of the average force of air resistance on the rocket during the upward trip?

__________ N

17.3 On an airless moon, you drop a golf ball of mass 106 g out of a skyscraper window. After it has fallen 125 meters, it is moving at 19.0 m/s. What is the rate of freefall acceleration on this moon?

__________ m/s^2

17.4 You are pulling your sister on a sled to the top of a 16.0 m high, frictionless hill with a 10.0° incline. Your sister and the sled have a total mass of 50.0 kg. You pull the sled, starting from rest, with a constant force of 127 N at an angle of 45.0° to the hill. If you pull from the bottom to the top, what will the speed of the sled be when you reach the top?

__________ m/s

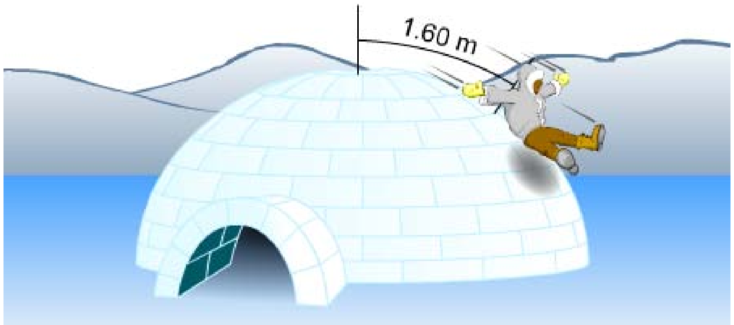
Section 19 - Conservation of energy

19.1 A large block of ice is moving down the hill toward you at 25.0 m/s. Its mass is 125 kg. It is sliding down a slope that makes a 30.0 degree angle with the horizontal. In short: Think avalanche. Assume the block started stationary and moves down the hill with zero friction. How many meters has it been sliding?

__________ m

19.2 Jin is sitting on top of a hemispherical, frictionless igloo of radius 2.40 meters. His friend pushes him, giving him an initial speed. Jin slides along the igloo and loses contact with it after he has traveled 1.60 meters along the surface. What was his initial speed?

__________ m/s



Section 20 - Sample problem: conservation of energy

20.1 A spring has its right end fixed and is installed on a horizontal table so that the free end, in equilibrium, is at x = 3.00 m. A 1.65 kg block coming from the left slides along the table. When it passes the origin, it is moving at 5.58 m/s. It strikes the spring, compresses it momentarily, and is then sent back toward the left, where it eventually comes to rest at the point x = 1.50 m. The coefficient of kinetic friction between the block and the table is 0.300. By what distance was the spring compressed?

__________ m

20.2 A 75.0 kg mass sits on an inclined plane, and a rope passing over a pulley at the top connects it to a hanging 125 kg mass. The pulley is frictionless and its mass is negligible. The coefficient of friction between the 75.0 kg block and the plane is 0.143. The system is released from rest, and after dropping 8.75 m, the 125 kg mass is moving at a speed of 8.00 m/s. What is the angle of inclination of the plane from the horizontal?

__________ °

Section 22 - Interactive problem: conservation of energy

22.1 Use the information given in the interactive problem in this section to answer the following question. What initial height is required for the soapbox car to make it through the hoop? Test your answer using the simulation.

__________ m

Section 24 - Interactive problem: a non-conservative force

24.1 Use the information given in the interactive problem in this section to answer the following question. What is the initial speed required for the car to overcome friction and jump through the hoop? Test your answer using the simulation.

__________ m/s

Section 25 - Interactive problem: a non-conservative force, part II

25.1 Use the information given in the interactive problem in this section to answer the following question. What height for the car will allow the car to jump through the hoop? Give your answer with 2 significant figures.

__________ m

Section 27 - Interactive summary problem: work, energy and power

27.1 The brakes in the go-kart from the simulation in the interactive problem in this section have broken. They now have only two states: on, which applies a braking force of 790 N, or off. Will you be able to make it through the course without crashing into the hay bales?

○ Yes ○ No

27.2 Using the information given in the interactive problem in this section, if you decide to apply 7200 W of power, what should the braking force be to stop the go-kart before the hay bales? Give your answer with 2 significant figures.

__________ N

27.3 You have decided to try to beat your personal best speed in your go-kart of 12 m/s. (a) Using the simulation in the interactive problem in this section, if you want to reach a speed of 13 m/s at the top of the loop-the-loop, how much power must you supply to the go-kart? (b) What braking force will you need to apply in order to stop the go-kart before you hit the hay bales?

(a) __________ W
(b) __________ N

Additional Problems

A.1 Fritz Strobl thrilled the world when he won the gold medal in the Salt Lake City games of 2002 in a daring run down an alpine skiing course. The course had a vertical drop of 880 meters. Assume his highest speed was 140 km/h, and that he was moving at that speed at the end. (a) How fast would he have been moving if he could have "ignored" forces like air resistance and friction? (b) How much energy did he lose to forces like air resistance, friction, and so forth (assume his mass is 80 kg, and express the answer as a positive number)?

(a) __________ m/s
(b) __________ J

A.2 A cannon fires a cannonball of mass 16.0 kg by applying a force of 2750 N along the 1.25 m length of the barrel. (a) How much work does the cannon do on the cannonball? (b) The cannon is aimed at a 25.0° angle above the horizontal. Assume gravity is the only other force acting on the cannonball as it moves through the cannon barrel. (That is, ignore all frictional forces.) What is the net work done by these 2 forces on the cannonball while it is in the cannon barrel?

(a) __________ J
(b) __________ J

A.3 When playing shuffleboard, a player exerts a constant force of 2.1 N on an initially stationary puck, at an angle 55° below the horizontal. If the player pushes the puck for 1.5 m, how fast is the puck moving when it is released? The mass of a puck is 0.49 kg. Ignore the force of friction.

__________ m/s

A.4 An elevator car, with a mass of 500 kg, starts at rest and moves upward with a constant acceleration until it reaches a velocity of 1.60 m/s after 2.70 seconds. If the car rises 2.16 m during this time, what is the average power of the elevator motor during this period? Ignore friction and other resistive forces, and the mass of the cables.

__________ W

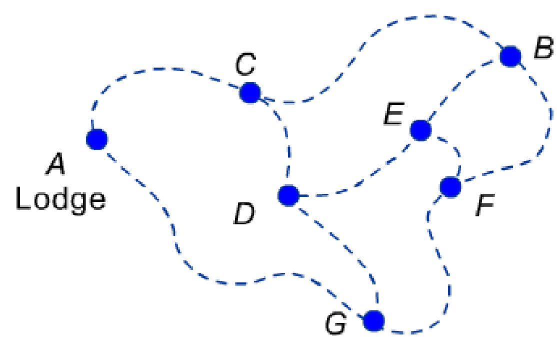
A.5 A goalie for the newly-created sport of "ice soccer" exerts 10.0 N of force opposite to the velocity of a soccer ball skidding along the ice without rolling, over a distance of 0.8 m. The ball comes to a complete stop. How fast must the 0.500 kg ball have been traveling when the goalie's hands first made contact with it? Ignore friction.

__________ m/s

A.6 You have a 0.750 kg picture frame on a mantle 1.30 m high. How much work was required to lift the frame from the ground to the mantle?

__________ J

A.7 The diagram shows a map of a local cross country ski area. You are starting at the lodge, marked as point A on the map, and want to go to point B. How much work is required to get to point B given the following information:
1) To go from D to E requires −300 J of work.
2) To go from D to G requires 200 J of work.
3) To go from A to G requires 400 J of work.
4) To go from B to E requires −200 J of work.



__________ J

A.8 When your sled starts down from the top of a hill, it hits a frictionless ice slick that extends all the way down the hill. At the bottom, the ground is dry and level. The effective coefficient of friction between the sled runners and the ground is 0.62. If the hill is 50 m high, how far will your sled travel once it reaches the bottom?

__________ m

A.9 The Queen Mary 2, whose maiden voyage was in January 2004, is a cruise ship that has a mass of 150,000 gross tons (which equals 1.52×10^8 kg, about three times that of the Titanic). Her electrically driven pod motors have a maximum power rating of 1.57×10^5 hp, or 117 MW. (a) What is the kinetic energy of the QM2 when she is moving at 15.0 meters/second? (b) Find the absolute minimum time in which the ship's engines could accelerate her from rest up to 15.0 m/s. Ignore the drag resistance of the water, air, and so on. (c) What is the force that the ship's propellers exert on the water when the Queen Mary 2 is moving at 15.0 m/s (assume that the maximum power is used)?

(a) __________ J
(b) __________ s
(c) __________ N

A.10 Lance Armstrong bikes at a constant speed up the *Alpe d'Huez*, a famous mountain pass. Assume his teammates do such a good job riding ahead of him that he can draft behind them and encounter no air resistance. This climb is described as "beyond classification" in terms of its difficulty. The climb is 13.8 km long at a 7.9% average grade (the grade, as a decimal, is the tangent of the angle of inclination). Assume that the combined mass of Lance and his bicycle is 83 kilograms. What is the magnitude of the work he does against the force of gravity?

__________ J

A.11 A block with mass m = 4.00 kg is attached to a spring with spring constant k = 500 N/m, and it slides on a horizontal table that is not frictionless. At a time when the spring is compressed by 3.00 cm, the block is observed to be moving toward the right (decompressing the spring) at 0.600 m/s. When the block reaches the point where the spring is completely relaxed, the block is measured to be moving at 0.550 m/s. (a) What is the net work done on the block during the time interval? (b) What is the work done by the spring? (c) What is the work done by friction? (d) What is the coefficient of kinetic friction between the block and the table?

(a) __________ J
(b) __________ J
(c) __________ J
(d) __________

A.12 Lamborghini states that its 2004 Murciélago® has a mass of 1650 kg. On a particular test run, its 580 hp (433 kW) engine accelerates the car from 0 to 100 km/h (62 mph) in 3.60 seconds. Assume the engine is working at its maximum power. How much energy is consumed by dissipative forces like air resistance and friction as the car accelerates from 0 to 100 km/h?

__________ J

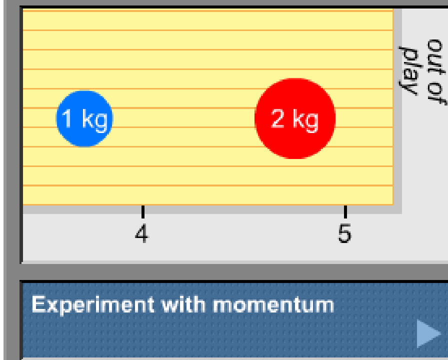
A.13 A motor lifts a 70.0 kg box off the ground, starting from rest. In 8.00 seconds it lifts the box to a height of 20.0 m. At that time, the box is moving upward with a velocity of 5.00 m/s. What is the average power of the motor during this time interval?

__________ W

8 Momentum

8.0 - Introduction

interactive 1



“The more things change, the more they stay the same” is a well-known French saying.

However, though witty and perhaps true for many matters on which the French have great expertise, this saying is simply not good physics.

Instead, a physicist would say: “Things stay the same, period. That is, unless acted upon by a net force.” Perhaps a little less *joie de vivre* than your average Frenchman, but nonetheless the key to understanding momentum.

What we now call momentum, Newton referred to as “quantity of motion.” The linear momentum of an object equals the product of its mass and velocity. (In this chapter, we focus on linear momentum. Angular momentum, or momentum due to rotation, is a topic in another chapter.) Momentum is a useful concept when applied to collisions, a subject that can be a lot of fun. In a collision, two or more objects exert forces on each other for a brief instant of time, and these forces are significantly greater than any other forces they may experience during the collision.

At the right is a simulation − a variation of shuffleboard − that you can use to begin your study of momentum and collisions. You can set the initial velocity for both the blue and the red pucks and use these velocity settings to cause them to collide. The blue puck has a mass of 1.0 kg, and the red puck a mass of 2.0 kg. The shuffleboard has no friction, but the pucks stop moving when they fall off the edge. Their momenta and velocities are displayed in output gauges.

Using the simulation, answer these questions. First, is it possible to have negative momentum? If so, how can you achieve it? Second, does the collision of the pucks affect the sum of their velocities? In other words, does the sum of their velocities remain constant? Third, does the collision affect the sum of their momenta? Remember to consider positive and negative signs when summing these values. Press PAUSE before and after the collisions so you can read the necessary data. For an optional challenge: Does the collision conserve the total kinetic energy of the pucks? If so, the collision is called an elastic collision. If it reduces the kinetic energy, the collision is called an inelastic collision.

8.1 - Momentum

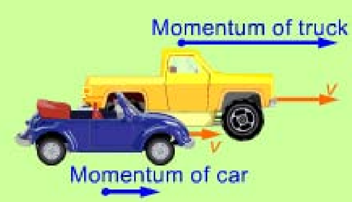
Momentum (linear): Mass times velocity.

An object’s linear momentum equals the product of its mass and its velocity. A fast moving locomotive has greater momentum than a slowly moving ping-pong ball.

The units for momentum are kilogram·meters/second (kg·m/s). A ping-pong ball with a mass of 2.5 grams moving at 1.0 m/s has a momentum of 0.0025 kg·m/s. A 100,000 kg locomotive moving at 5 m/s has a momentum of 5×10^5 kg·m/s.

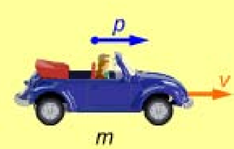
Momentum is a vector quantity. The momentum vector points in the same direction as the velocity vector. This means that if two identical locomotives are moving at the same speed and one is heading east and the other west, they will have equal but **opposite** momenta, since they have equal but oppositely directed velocities.

concept 1



Momentum
Moving objects have momentum
Momentum increases with mass, velocity

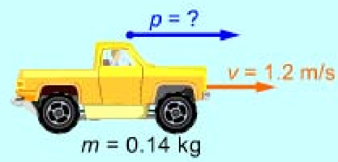
equation 1



$$\mathbf{p} = m\mathbf{v}$$

p = momentum
m = mass
v = velocity
Momentum same direction as velocity
Units: kg·m/s

example 1



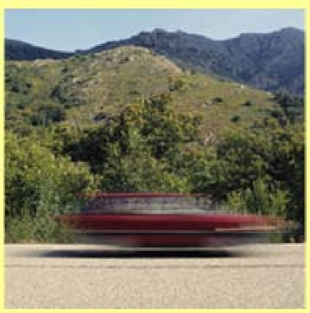
What is the toy truck's momentum?

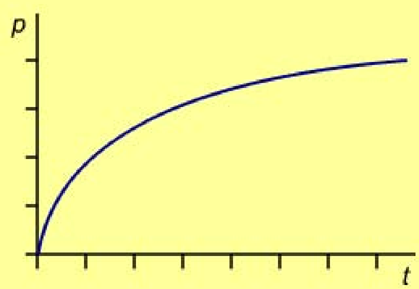
$p = mv$

p = (0.14 kg)(1.2 m/s)

p = 0.17 kg·m/s to the right

8.2 - Momentum and Newton's second law





equation 1

Momentum and Newton's second law

$$\Sigma \mathbf{F} = \frac{\Delta \mathbf{p}}{\Delta t}$$

Σ**F** = net force
p = momentum
t = time

Although you started your study of physics with velocity and acceleration, early physicists such as Newton focused much of their attention on momentum.

The equation in Equation 1 may remind you of Newton's second law. In fact, it is equivalent to the second law (as we show below) and Newton stated his law in this form. This equation is useful when you know the change in the momentum of an object (or, equivalently for an object of constant mass, the change in velocity). This equation as stated assumes an average or constant external net force.

Below, we show that this equation is equivalent to the more familiar version of Newton's second law based on mass and acceleration. We state that version of Newton's law, and then use the definition to restate the law as you see it here.

Step	Reason
1. $\Sigma \mathbf{F} = m\mathbf{a} = m\frac{\Delta \mathbf{v}}{\Delta t}$	Newton's second law; definition of acceleration
2. $\mathbf{p} = m\mathbf{v}$	definition of momentum
3. $\frac{\Delta \mathbf{p}}{\Delta t} = m\frac{\Delta \mathbf{v}}{\Delta t}$	divide equation 2 by Δt
4. $\Sigma \mathbf{F} = \frac{\Delta \mathbf{p}}{\Delta t}$	substitute equation 3 into equation 1

concept 1

Impulse
Force times elapsed time
Change in momentum

Impulse: Change in momentum.

In the prior section, we stated that net force equals change in momentum per unit time. We can rearrange this equation and state that the change in momentum equals the product of the average force and the elapsed time, which is shown in Equation 1.

The change in momentum is called the impulse of the force, and is represented by $\mathbf{J}$. Impulse is a vector, has the same units as momentum, and points in the same direction as the change in momentum and as the force. The relationship shown is called the *impulse-momentum theorem*.

In this section, we focus on the case when a force is applied for a brief interval of time, and is stated or approximated as an average force. This is a common and important way of applying the concept of impulse.

A net force is required to accelerate an object, changing its velocity and its momentum. The greater the net force, or the longer the interval of time it is applied, the more the object's momentum changes, which is the same as saying the impulse increases.

Engineers apply this concept to the systems they design. For instance, a cannon barrel is long so that the cannonball is exposed to the force of the explosive charge longer, which causes the cannonball to experience a greater impulse, and a greater change in momentum.

Even though a longer barrel allows the force to be applied for a longer time interval, it is still brief. Measuring a rapidly changing force over such an interval may be difficult, so the force is often modeled as an average force. For example, when a baseball player swings a bat and hits the ball, the duration of the collision can be as short as 1/1000th of a second and the force averages in the thousands of newtons.

The brief but large force that the bat exerts on the ball is called an *impulsive force*. When analyzing a collision like this, we ignore other forces (like gravity) that are acting upon the ball because their effect is minimal during this brief period of time.

In the illustration for Equation 1, you see a force that varies with time (the curve) and the average of that force (the straight dashed line). The area under the curve and the area of the rectangle both equal the impulse, since both equal the product of force and time.

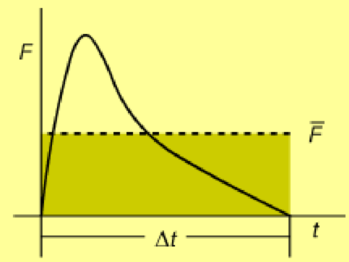
The nature of impulse explains why coaches teach athletes like long jumpers, cyclists, skiers and martial artists to relax when they land or fall, and why padded mats and sand pits are used. In Example 1 on the right, we calculate the (one-dimensional) impulse experienced by a long jumper on landing in the sand pit, from her change in momentum.

Her impulse (change in momentum) is the same, however long it takes her to stop when she hits the ground. In the example, that impulse is −530 kg·m/s.

Why does she want to extend the time of her landing? She wants to make this time as long as possible (by landing in the sand, by flexing her knees), since it means the collision lasts longer. Since impulse equals force multiplied by elapsed time, the average force required to produce the change in momentum **decreases** as the time **increases**. The reduced average force lessens the chance of injury. Padded mats are another application of this concept: The impulse of landing is the same on a padded or unpadded floor, but a mat increases the duration of a landing and reduces its average force.

There are also numerous applications of this principle outside of sports. For example, cars have “crumple zones” designed into them that collapse upon impact, extending the duration of the impulse during a collision and reducing the average force.

equation 1

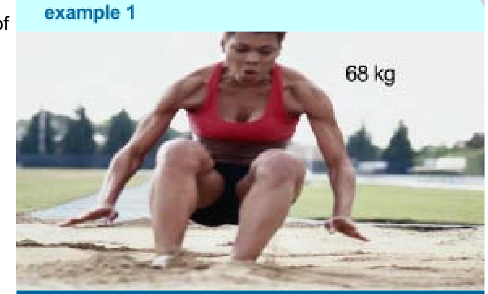


Impulse

$$\mathbf{J} = \bar{\mathbf{F}}\Delta t = \Delta \mathbf{p}$$

$\mathbf{J}$ = impulse
$\bar{\mathbf{F}}$ = average force
Δt = elapsed time
$\mathbf{p}$ = momentum
Units of impulse: kg·m/s

example 1



The long jumper's speed just before landing is 7.8 m/s. What is the impulse of her landing?

$\mathbf{J} = \mathbf{p}_f - \mathbf{p}_i$

$J = mv_f - mv_i$

$J = 0.0 - (68 \text{ kg})(7.8 \text{ m/s})$

$J = -530 \text{ kg·m/s}$

8.4 - Physics at play: hitting a baseball

When a professional baseball player swings a bat and hits a ball square on, he will dramatically change its velocity in a millisecond. A fastball can approach the plate at around 95 miles per hour, and in a line drive shot, the ball can leave the bat in roughly the opposite direction at about 110 miles per hour, a change of about 200 mph in about a millisecond.

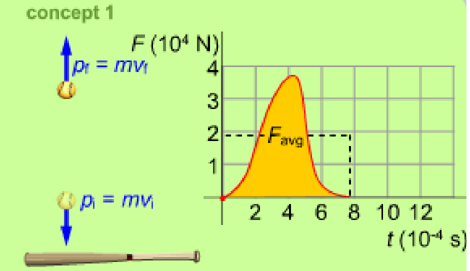
The bat exerts force on the baseball in the very brief period of time they are in contact. The amount of force varies over this brief interval, as the graph to the right reflects.

At the moment of contact, the bat and ball are moving toward each other. The force on the ball increases as they come together and the ball compresses against the bat. The force applied to the ball during the time it is in contact with the bat is responsible for the ball's change in momentum.

How long the bat stays in contact with the ball is much easier to measure than the average force the bat exerts on the ball, but by applying the concept of impulse, that force can be calculated. Impulse equals both the average force times the elapsed time and the change in momentum. Since the velocities of the baseball can be observed (say, with a radar gun), and the baseball's mass is known, its change in momentum can be calculated, as we do in Example 1. The time of the collision can be observed using stroboscopic photography and other techniques. This leaves one variable − average force − and we solve for that in the example problem.

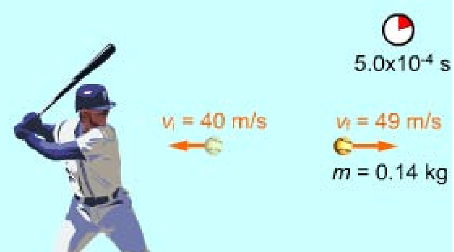
The average force equals 2.5×10^4 N. A barrier that stops a car moving at 20 miles per hour in half a second exerts a comparable average amount of force.

concept 1



Baseballs, bats and impulse
Force applied over time changes momentum
Impulse = change in momentum
Impulse = average force × elapsed time

example 1



The ball arrives at 40 m/s and leaves at 49 m/s in the opposite direction. The contact time is 5.0×10^{-4} s. What is the average force on the ball?

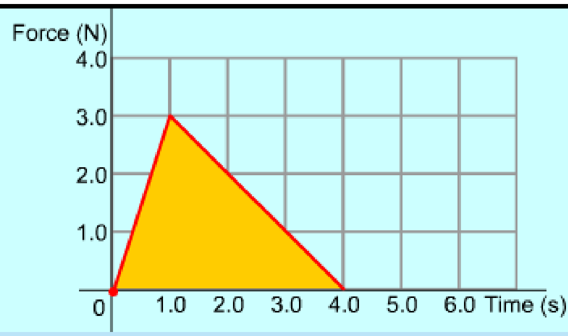
$J = F_{avg}\,\Delta t = \Delta p = m\Delta v$

$F_{avg} = m\Delta v/\Delta t$

$$F_{avg} = \frac{(0.14\ \text{kg})(49 - (-40)\ \text{m/s})}{5.0 \times 10^{-4}\ \text{s}}$$

$F_{avg} = 2.5 \times 10^4$ N

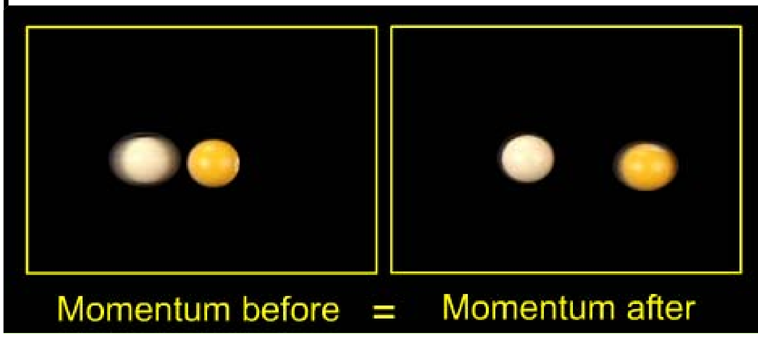
8.5 - Interactive checkpoint: force graph and impulse



An object has force applied to it in a constant direction. The graph shows how the force varies with time. If the object has a mass of 5.0 kg, what is its change in velocity over the time the force is applied?

Answer:

$\Delta v =$ ______ m/s



concept 1

Conservation of momentum
No net external force on system:
Total momentum is conserved

Conservation of momentum: The total momentum of an isolated system is constant.

Momentum is conserved in an isolated system. An *isolated system* is one that does not interact with its environment. Momentum can transfer from object to object within this system but the vector sum of the momenta of all the objects remains constant.

Excluding deep space, it can be difficult to find locations where a system has no interaction with its environment. This makes it useful to state that momentum is conserved in a system that has no net force acting on it.

We will use pool balls on a pool table to discuss the conservation of momentum. To put it more formally, the pool balls are the system.

In pool, a player begins play by striking the white cue ball. To use the language of physics, a player causes there to be a net external force acting on the cue ball.

Once the cue ball has been struck, it may collide with another ball, and more collisions may ensue. However, there is **no** net external force acting on the balls after the cue ball has been struck (ignoring friction which we will treat as negligible). The normal force of the table balances the force of gravity on the balls. Since there is now no net external force acting on the balls, the total amount of their momentum remains constant. When they collide, the balls exert forces on one another, but this is a force internal to the system, and does not change its total momentum.

In the scenario you see illustrated above, the cue ball and another ball are shown before and after a collision. The cue ball initially has positive momentum since it is moving to the right. The ball it is aimed at is initially stationary and has zero momentum.

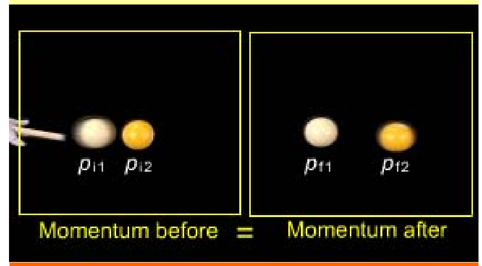
When the cue ball strikes its target, the cue ball slows down and the other ball speeds up. In fact, the cue ball may stop moving. You see this shown on the right side of the illustration above. During the collision, momentum transfers from one ball to another. The law of conservation of momentum states that the combined momentum of both remains constant: One ball's loss equals the other ball's gain.

A rifle also provides a notable example of the conservation of momentum. Before it is fired, the initial momentum of a rifle and the bullet it fires are both zero (since neither has any velocity). When the rifle is fired, the bullet moves in one direction and the rifle recoils in the opposite direction. The bullet and the rifle each now have nonzero momentum, but the vector sum of their momenta must remain at zero.

Two factors account for this. First, the rifle and the bullet are moving in opposite directions. In the case of the rifle and bullet, all the motion takes place along a line, so we can use positive and negative to indicate direction. Let's assume the bullet has positive velocity; since the rifle moves (recoils) in the opposite direction, it has negative velocity. The momentum vector of each object points in the same direction as its velocity vector. This means the bullet has positive momentum while the rifle has negative momentum.

Second, for momentum to be conserved, the sum of these momenta must equal zero, since the sum was zero before the rifle was fired. The amount of momentum of the faster moving but less massive bullet equals the amount of momentum of the more massive but slower moving rifle. When the two are added together, the total momentum continues to equal zero.

equation 1

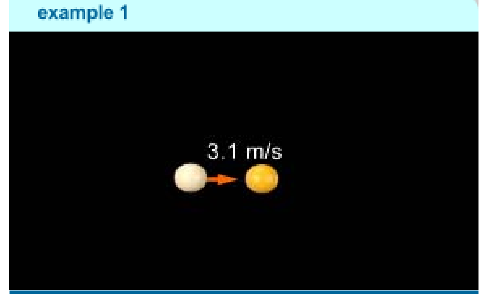


Conservation of momentum

$$\boxed{\mathbf{p}_{i1} + \mathbf{p}_{i2} + \ldots + \mathbf{p}_{in} = \mathbf{p}_{f1} + \mathbf{p}_{f2} + \ldots + \mathbf{p}_{fn}}$$

$\mathbf{p}_{i1}, \mathbf{p}_{i2}, \ldots, \mathbf{p}_{in}$ = initial momenta
$\mathbf{p}_{f1}, \mathbf{p}_{f2}, \ldots, \mathbf{p}_{fn}$ = final momenta

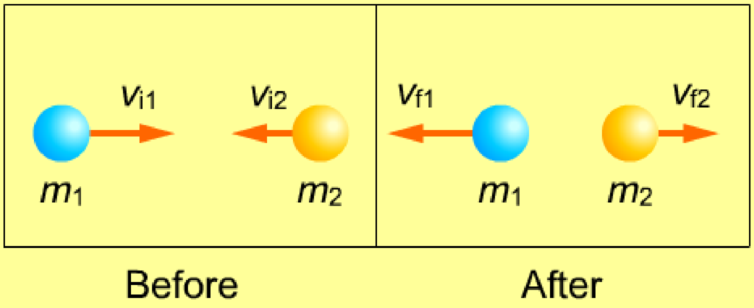
example 1



The balls have the same mass. The cue ball strikes the stationary yellow ball head on, and stops. What is the yellow ball's resulting velocity?

$\mathbf{p}_{i,cue} + \mathbf{p}_{i,yel} = \mathbf{p}_{f,cue} + \mathbf{p}_{f,yel}$
$m\mathbf{v}_{i,cue} + m\mathbf{v}_{i,yel} = m\mathbf{v}_{f,cue} + m\mathbf{v}_{f,yel}$
$\mathbf{v}_{i,cue} + \mathbf{v}_{i,yel} = \mathbf{v}_{f,cue} + \mathbf{v}_{f,yel}$
$v_{i,cue} + v_{i,yel} = v_{f,cue} + v_{f,yel}$
3.1 m/s + 0.0 m/s = 0.0 m/s + $v_{f,yel}$
$\mathbf{v}_{f,yel}$ = 3.1 m/s to the right

8.7 - Derivation: conservation of momentum from Newton's laws



equation 1

Conservation of momentum

$$m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2} = m_1\mathbf{v}_{f1} + m_2\mathbf{v}_{f2}$$

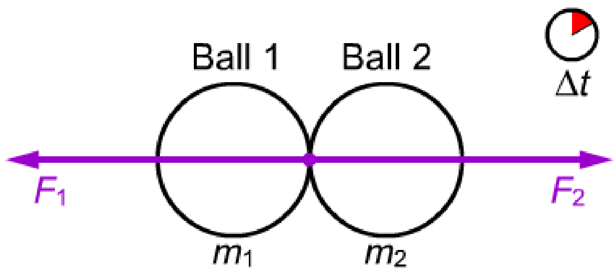
$\mathbf{p} = m\mathbf{v}$ = momentum
m_1, m_2 = masses of objects
$\mathbf{v}_{i1}, \mathbf{v}_{i2}$ = initial velocities
$\mathbf{v}_{f1}, \mathbf{v}_{f2}$ = final velocities

Newton formulated many of his laws concerning motion using the concept of momentum, although today his laws are stated in terms of force and acceleration. The law of conservation of momentum can be derived from his second and third laws.

The derivation uses a collision between two balls of masses m_1 and m_2 with velocities v_1 and v_2. You see the collision illustrated above, along with the conservation of momentum equation we will prove for this situation.

To derive the equation, we consider the forces on the balls during their collision. During the time Δt of the collision, the balls exert forces $\mathbf{F}_1$ and $\mathbf{F}_2$ on each other. ($\mathbf{F}_1$ is the force on ball 1 and $\mathbf{F}_2$ the force on ball 2.)

Diagram



This diagram shows the forces on the balls during the collision.

Variables

duration of collision	Δt	
	ball 1	ball 2
force on ball	$\mathbf{F}_1$	$\mathbf{F}_2$
mass	m_1	m_2
acceleration	$\mathbf{a}_1$	$\mathbf{a}_2$
initial velocity	$\mathbf{v}_{i1}$	$\mathbf{v}_{i2}$
final velocity	$\mathbf{v}_{f1}$	$\mathbf{v}_{f2}$

Strategy

1. Use Newton's third law: The forces will be equal but opposite.
2. Use Newton's second law, $\mathbf{F} = m\mathbf{a}$, to determine the acceleration of the balls.
3. Use the definition that expresses acceleration in terms of change in velocity. This will result in an equation that contains momentum ($m\mathbf{v}$) terms.

Physics principles and equations

In addition to Newton's laws cited above, we will use the definition of acceleration.

$$\mathbf{a} = \Delta\mathbf{v}/\Delta t$$

Step-by-step derivation

In these first steps, we use Newton's third law followed by his second law.

Step	Reason
1. $\mathbf{F}_1 = -\mathbf{F}_2$	Newton's third law
2. $m_1\mathbf{a}_1 = -m_2\mathbf{a}_2$	Newton's second law
3. $m_1\dfrac{\Delta\mathbf{v}_1}{\Delta t} = -m_2\dfrac{\Delta\mathbf{v}_2}{\Delta t}$	definition of acceleration

In the next steps, we apply the definition of the change Δv in velocity. After some algebraic simplification we obtain the result we want: The sum of the initial momenta equals the sum of the final momenta.

Step	Reason
4. $m_1\dfrac{\mathbf{v}_{f1} - \mathbf{v}_{i1}}{\Delta t} = -m_2\dfrac{\mathbf{v}_{f2} - \mathbf{v}_{i2}}{\Delta t}$	definition of change in velocity
5. $m_1\mathbf{v}_{f1} - m_1\mathbf{v}_{i1} = -m_2\mathbf{v}_{f2} + m_2\mathbf{v}_{i2}$	multiply both sides by Δt
6. $m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2} = m_1\mathbf{v}_{f1} + m_2\mathbf{v}_{f2}$	rearrange

8.8 - Interactive checkpoint: astronaut

The 55.0 kg astronaut is stationary in the spaceship's reference frame. She wants to move at 0.500 m/s to the left. She is holding a 4.00 kg bag of dehydrated astronaut chow. At what velocity must she throw the bag to achieve her desired velocity? (Assume the positive direction is to the right.)

Answer:

v_{fb} = ________ m/s

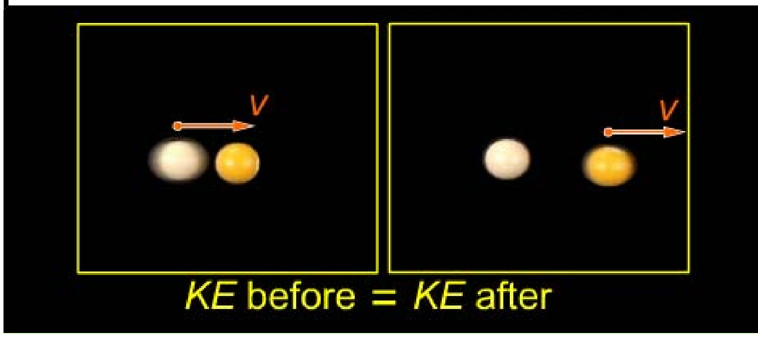
8.9 - Interactive checkpoint: bumper cars

Hal sneaks up behind Larry, whose bumper car is at rest, and bumps his car. Hal's car bounces off Larry's and heads backward. Larry's car is propelled forward at 0.800 m/s. Hal's velocity before the collision is 1.50 m/s. What is Hal's velocity after the collision?

Answer:

v_{fH} = ________ m/s

8.10 - Collisions



concept 1

Elastic collision
Kinetic energy is conserved

Elastic collision: The kinetic energy of the system is unchanged by the collision.

Inelastic collision: The kinetic energy of the system is changed by the collision.

In a collision, one moving object briefly strikes another object. During the collision, the forces the objects exert on each other are much greater than the net effect of other forces acting on them, so we may ignore these other forces.

Elastic and inelastic are two terms used to define types of collisions. These types of collisions differ in whether the total amount of **kinetic** energy in the system stays constant or is reduced by the collision. In any collision, the system's total amount of energy must be the same before and after, because the law of conservation of energy must be obeyed. But in an inelastic collision, some of the kinetic energy is transformed by the collision into other types of energy, so the total kinetic energy decreases.

For example, a car crash often results in dents. This means some kinetic energy compresses the car permanently; other KE becomes thermal energy, sound energy and so on. This means that an **inelastic** collision reduces the total amount of KE.

In contrast, the total kinetic energy is the same before and after an **elastic** collision. None of the kinetic energy is transformed into other forms of energy. The game of pool provides a good example of nearly elastic collisions. The collisions between balls are almost completely elastic and little kinetic energy is lost when they collide.

In both elastic and inelastic collisions occurring within an isolated system, momentum is conserved. This important principle enables you to analyze any collision.

We will mention a third type of collision briefly here: *explosive collisions*, such as what occurs when a bomb explodes. In this type of collision, the kinetic energy is greater after the collision than before. However, since momentum is conserved, the explosion does not change the total momentum of the constituents of the bomb.

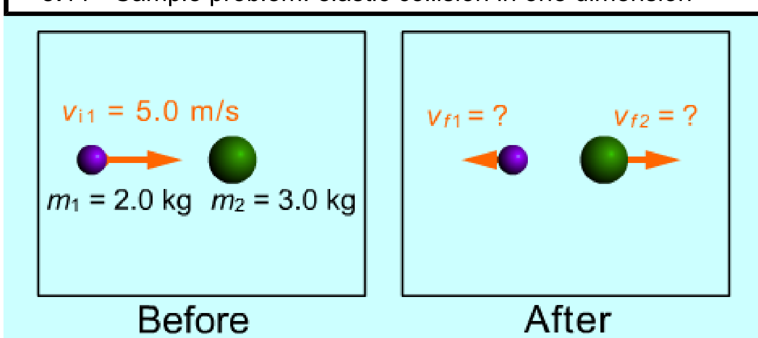
concept 2

Inelastic collision
Kinetic energy is not conserved

concept 3

Either type of collision
Momentum is conserved

8.11 - Sample problem: elastic collision in one dimension



The small purple ball strikes the stationary green ball in an elastic collision. What are the final velocities of the two balls?

The picture and text above pose a classic physics problem. Two balls collide in an elastic collision. The balls collide head on, so the second ball moves away along the same line as the path of the first ball. The balls' masses and initial velocities are given. You are asked to calculate their velocities after the collision. The strategy for solving this problem relies on the fact that both the momentum and kinetic energy remain unchanged.

Variables

	ball 1 (purple)	ball 2 (green)
mass	$m_1 = 2.0$ kg	$m_2 = 3.0$ kg
initial velocity	$v_{i1} = 5.0$ m/s	$v_{i2} = 0$ m/s
final velocity	v_{f1}	v_{f2}

What is the strategy?

1. Set the momentum before the collision equal to the momentum after the collision.
2. Set the kinetic energy before the collision equal to the kinetic energy after the collision.
3. Use algebra to solve two equations with two unknowns.

Physics principles and equations

Since problems like this one often ask for values **after** a collision, it is convenient to state the following conservation equations with the final values on the left.

Conservation of momentum

$$m_1\mathbf{v}_{f1} + m_2\mathbf{v}_{f2} = m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2}$$

Conservation of kinetic energy

$$½\, m_1 v_{f1}{}^2 + ½\, m_2 v_{f2}{}^2 = ½\, m_1 v_{i1}{}^2 + ½\, m_2 v_{i2}{}^2$$

Step-by-step solution

First, we use the conservation of momentum to find an equation where the only unknown values are the two final velocities. Since all the motion takes place on a horizontal line, we use sign to indicate direction.

Step	Reason
1. $m_1 v_{f1} + m_2 v_{f2} = m_1 v_{i1} + m_2 v_{i2}$	conservation of momentum
2. $(2.0\text{ kg})v_{f1} + (3.0\text{ kg})v_{f2}$ $= (2.0\text{ kg})(5.0\text{ m/s}) + (3.0\text{ kg})(0.0\text{ m/s})$	enter values
3. $v_{f1} = (-1.5)v_{f2} + 5.0$	solve for v_{f1}

The conservation of kinetic energy gives us another equation with these two unknowns.

Step	Reason
4. $\frac{1}{2}m_1 v_{f1}^2 + \frac{1}{2}m_2 v_{f2}^2 = \frac{1}{2}m_1 v_{i1}^2 + \frac{1}{2}m_2 v_{i2}^2$	elastic collision: *KE* conserved
5. $m_1 v_{f1}^2 + m_2 v_{f2}^2 = m_1 v_{i1}^2 + m_2 v_{i2}^2$	simplify
6. $(2.0\text{ kg})v_{f1}^2 + (3.0\text{ kg})v_{f2}^2$ $= (2.0\text{ kg})(5.0\text{ m/s})^2 + (3.0\text{ kg})(0.0\text{ m/s})^2$	enter values
7. $(2.0)v_{f1}^2 + (3.0)v_{f2}^2 - 50.0 = 0$	re-arrange as quadratic equation

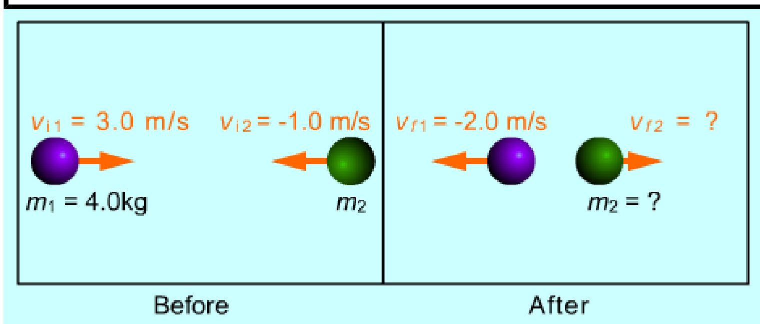
We substitute the expression for the first ball's final velocity found in equation 3 into the quadratic equation, and solve. This gives us the second ball's final velocity. Then we use equation 3 again to find the first ball's final velocity. One velocity is negative, and one positive − one ball moves to the left after the collision, the other to the right.

Step	Reason
8. $(2.0)[(-1.5)v_{f2}+(5.0)]^2+(3.0)v_{f2}^2-50=0$	substitute equation 4 into equation 8
9. $15v_{f2}^2-60v_{f2}=0$	simplify
10. $v_{f2}=4.0$ m/s (to the right)	solve equation
11. $v_{f1}=(-1.5)v_{f2}+5.0$ $v_{f1}=-1.0$ m/s (to the left)	use equation 3 to find v_{f1}

A quick check shows that the total momentum both before and after the collision is 10 kg·m/s. The kinetic energy is 25 J in both cases. This verifies that we did the computations correctly.

There is a second solution to this problem: You can see that $v_{f2}=0$ is also a solution to the quadratic equation in step 9, and then using step 3, you see that v_{f1} would equal 5.0 m/s. This solution satisfies the conditions that momentum and KE are conserved, and it describes what happens if the balls do not collide. In other words, the purple ball passes by the green ball without striking it.

8.12 - Interactive checkpoint: another one dimensional collision problem



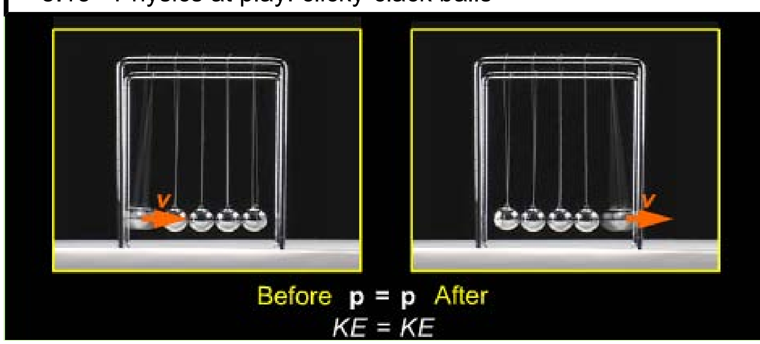
Two balls move toward each other and collide head-on in an elastic collision. What is the mass and final velocity of the green ball?

Answer:

$m_2 =$ [ ] kg

$v_{f2} =$ [ ] m/s

8.13 - Physics at play: clicky-clack balls



concept 1

Clicky-clack balls
Momentum conserved
Elastic collision: kinetic energy conserved

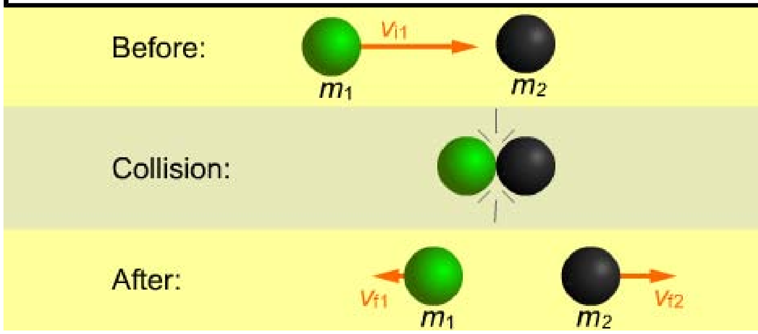
The law of conservation of momentum and the nature of elastic collisions underlie the functioning of a desktop toy: a set of balls of equal mass hanging from strings. This toy is shown in the photograph above.

You may have seen these toys in action but if you have not, imagine pulling back one ball and releasing it toward the pack of balls. (Click on Concept 1 to launch a video.) The one ball strikes the pack, stops, and a ball on the far side flies up, comes back down, and strikes the pack. The ball you initially released now flies back up again, returns to strike the pack, and so forth. The motion continues in this pattern for quite a while.

Interestingly, if you pick up two balls and release them, then two balls on the far side of the pack will fly off, resulting in a pattern of two balls moving. This pattern obeys the principle of the conservation of momentum as well as the definition of an elastic collision: kinetic energy remains constant. Other scenarios that on the surface might seem plausible fail to meet both criteria. For instance, if one ball moved off the pack at twice the speed of the two balls striking, momentum would be conserved, but the KE of the system would increase, since KE is a function of the **square** of velocity. Doubling the velocity of one ball quadruples its KE. One ball leaving at twice the velocity would have twice the combined KE of the two balls that struck the pack.

However, it turns out this is not the only solution which obeys the conservation of momentum and of kinetic energy. For instance, the striking ball could rebound at less than its initial speed, and the remaining four balls could move in the other direction as a group. With certain speeds for the rebounding ball and the pack of four balls, this would provide a solution that would obey both principles. Why the balls behave exactly as they do has inspired plenty of discussion in physics journals.

8.14 - Physics at play: elastic collisions and sports



equation 1

Elastic collision, m_2 initially stationary

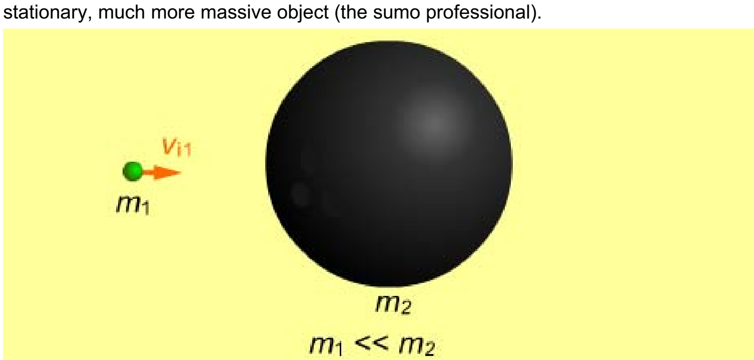
$$v_{f1} = \frac{m_1 - m_2}{m_1 + m_2} v_{i1}$$

$$v_{f2} = \frac{2m_1}{m_1 + m_2} v_{i1}$$

In previous sections, you saw that when a cue ball strikes a stationary pool ball of equal mass head on and stops, the second ball acquires all the momentum of the cue ball and all of its kinetic energy. This means the final velocity of the struck ball is the same as the initial velocity of the cue ball.

In this section, we will consider what happens in two other interesting scenarios for elastic collisions. We start with the two equations above. These equations hold true for head-on elastic collisions **when the second object is initially stationary**. They can be derived algebraically from the equations for conservation of momentum and kinetic energy, and we do so below.

But first, we consider what occurs in two situations when the objects have very different masses. The first is a hypothetical sumo-wrestling scenario in which we assume you are much less massive than the average sumo wrestler. Here, a much lighter object (you) runs into a stationary, much more massive object (the sumo professional).



equation 2

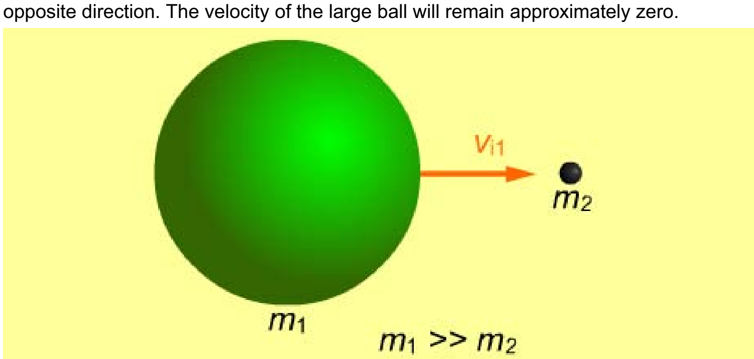
Moving object much less massive

$m_1 << m_2$, so m_1 treated as 0

$$v_{f1} = \frac{m_1 - m_2}{m_1 + m_2} v_{i1} \approx \frac{-m_2}{m_2} v_{i1} = -v_{i1}$$

$$v_{f2} = \frac{2m_1}{m_1 + m_2} v_{i1} \approx 0$$

When you run into the sumo wrestler, you bounce back with approximately the same speed at which you hit him, and he barely moves. The initial velocities are shown above, using two balls of differing mass to model the collision. As the equations show, the final velocity of the small moving ball will be approximately the negative of its initial velocity, which means that after the collision it moves at the same speed but in the opposite direction. The velocity of the large ball will remain approximately zero.



equation 3

Moving object much more massive

$m_1 >> m_2$, so m_2 treated as 0

$$v_{f1} = \frac{m_1 - m_2}{m_1 + m_2} v_{i1} \approx \frac{m_1}{m_1} v_{i1} = v_{i1}$$

$$v_{f2} = \frac{2m_1}{m_1 + m_2} v_{i1} \approx \frac{2m_1}{m_1} v_{i1} = 2v_{i1}$$

The other collision could be called "the quarterback sack." It is illustrated above. This occurs when a very massive object (like a lineman) collides elastically with a much less massive stationary object (the quarterback). The mathematics shows the result that the lineman continues with essentially the same velocity after the collision. The quarterback flies off at about twice the speed of the lineman in the same direction the lineman was moving. This result helps to explain the high concussion rate among quarterbacks, the quality of the football commentary of some ex-quarterbacks, and the increasing popularity of soccer.

In the steps that follow, we derive the two "final speed" equations that appear above.

$$v_{f1} = \frac{m_1 - m_2}{m_1 + m_2} v_{i1}$$

$$v_{f2} = \frac{2m_1}{m_1 + m_2} v_{i1}$$

Each of the equations gives the final speed of one of the objects in terms of the initial speed of the incoming object, multiplied by a quotient involving the masses of both objects. In both scenarios, **the initial velocity of the target object is zero**.

Variables

	incoming object	target object
mass	m_1	m_2
initial velocity	v_{i1}	$v_{i2} = 0$
final velocity	v_{f1}	v_{f2}

Strategy

1. Use the conservation of momentum and the conservation of energy to write "before and after" equations for the collision.
2. Solve for the final velocity of each ball by combining the two equations.

Physics principles and equations

Momentum is conserved in an isolated system

$$m_1 v_{i1} + m_2 v_{i2} = m_1 v_{f1} + m_2 v_{f2}$$

In an **elastic** collision, kinetic energy is conserved.

$$\tfrac{1}{2} m_1 v_{i1}^2 + \tfrac{1}{2} m_2 v_{i2}^2 = \tfrac{1}{2} m_1 v_{f1}^2 + \tfrac{1}{2} m_2 v_{f2}^2$$

Mathematics facts

We will use the difference of squares product formula.

$$(a - b)(a + b) = a^2 - b^2$$

Step-by-step derivation

We start in step 1 with the equations for conservation of momentum and kinetic energy. We manipulate these equations algebraically, making use of the difference of squares formula, to get the intermediate equation $v_{f2} = v_{i1} + v_{f1}$.

Step	Reason
1. (a) $m_1 v_{i1} = m_1 v_{f1} + m_2 v_{f2}$ (b) $\frac{1}{2} m_1 v_{i1}^2 = \frac{1}{2} m_1 v_{f1}^2 + \frac{1}{2} m_2 v_{f2}^2$	conservation of momentum (a) and energy (b)
2. (a) $m_1 (v_{i1} - v_{f1}) = m_2 v_{f2}$ (b) $m_1 (v_{i1}^2 - v_{f1}^2) = m_2 v_{f2}^2$	rearrange equations 1a and 1b
3. (a) $m_1 (v_{i1}^2 - v_{f1}^2) / v_{f2} = m_2 (v_{i1} + v_{f1})$ (b) $m_1 (v_{i1}^2 - v_{f1}^2) / v_{f2} = m_2 (v_{f2})$	multiply equation 2a by $(v_{i1} + v_{f1}) / v_{f2}$; divide equation 2b by v_{f2}
4. $v_{f2} = v_{i1} + v_{f1}$	set right sides of equations 3a and 3b equal and divide by m_2

Moving object. Next, we substitute the equation derived above back into the equation for conservation of momentum. We solve for the final velocity of the initially moving object.

Step	Reason
5. $m_1 v_{i1} = m_1 v_{f1} + m_2(v_{i1} + v_{f1})$	substitute equation 4 into equation 1a
6. $v_{f1} = \dfrac{m_1 - m_2}{m_1 + m_2} v_{i1}$	solve for v_{f1}

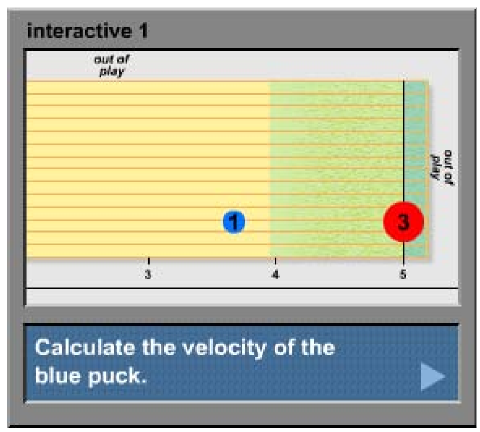
Target object. Now we combine equations 4 and 6 to get a sum that involves one fractional term. Using a common denominator to do the addition, we obtain an expression for the final velocity of the target object.

Step	Reason
7. $v_{f2} = v_{i1} + \dfrac{m_1 - m_2}{m_1 + m_2} v_{i1}$	substitute equation 6 into equation 4
8. $v_{f2} = \left[\dfrac{m_1 + m_2}{m_1 + m_2}\right] v_{i1} + \dfrac{m_1 - m_2}{m_1 + m_2} v_{i1}$	rewrite coefficient "1" as a fraction
9. $v_{f2} = \dfrac{2m_1}{m_1 + m_2} v_{i1}$	simplify

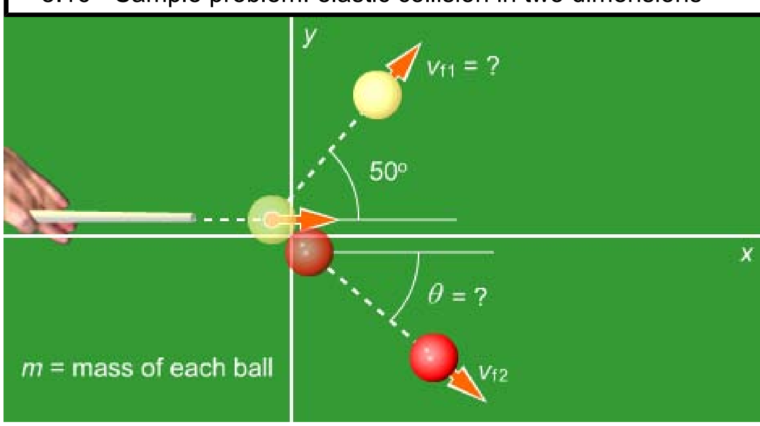
8.15 - Interactive problem: shuffleboard collisions

The simulation at the right shows a variation of the game of shuffleboard. The red puck has an initial velocity of −2.0 m/s. You want to set the initial velocity of the blue puck so that after the two pucks collide head on in an elastic collision, the red puck moves with a velocity of +2.0 m/s. This will cause the red puck to stop at the scoring line, since the friction in the green area on the right side of the surface will cause it to slow down and perhaps stop.

The blue puck has a mass of 1.0 kilograms. The red puck's mass is 3.0 kilograms. Use the fact that the collision is elastic to calculate and enter the initial velocity of the blue puck to the nearest 0.1 m/s and press GO to see the results. Press RESET to try again.



8.16 - Sample problem: elastic collision in two dimensions



The red ball is initially stationary. At what angle from the horizontal does it travel after the elastic collision?

Instead of striking the red ball head on, the cue ball above hits it off center. This means the motion takes place in both the x and y dimensions after the collision. You are asked to find the final angle of the red ball. Conservation of momentum still applies, but in two dimensions.

Variables

cue ball mass	m
red ball mass	m
cue ball initial velocity	$\mathbf{v}_{i1}$
cue ball final velocity	$\mathbf{v}_{f1}$
red ball initial velocity	$\mathbf{v}_{i2} = 0$ m/s
red ball final velocity	$\mathbf{v}_{f2}$
final angle of cue ball	50°
final angle of red ball	θ

Strategy

1. In an elastic collision like this one, kinetic energy is conserved.
2. Momentum is conserved. We write a vector equation that applies this principle.
3. Some vector arithmetic, and combining the two equations, allows us to calculate the desired angle.

Physics principles and equations

Conservation of momentum

$$m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2} = m_1\mathbf{v}_{f1} + m_2\mathbf{v}_{f2}$$

Conservation of kinetic energy

$$\frac{1}{2}m_1v_{i1}^2+\frac{1}{2}m_2v_{i2}^2=\frac{1}{2}m_1v_{f1}^2+\frac{1}{2}m_2v_{f2}^2$$

Mathematics principles

The dot product of two vectors, at an angle φ

$$\mathbf{A}\cdot\mathbf{B} = AB\cos\varphi$$

The dot product of a vector with itself has a simple result.

$$\mathbf{A}\cdot\mathbf{A} = A^2$$

Step-by-step solution

We begin with the conservation of kinetic energy, since this is an elastic collision.

Step	Reason
1. $\frac{1}{2}mv_{i1}^2+\frac{1}{2}mv_{i2}^2=\frac{1}{2}mv_{f1}^2+\frac{1}{2}mv_{f2}^2$	conservation of energy
2. $v_{i1}^2 = v_{f1}^2 + v_{f2}^2$	simplify; $v_{i2} = 0$

Momentum is conserved in any collision.

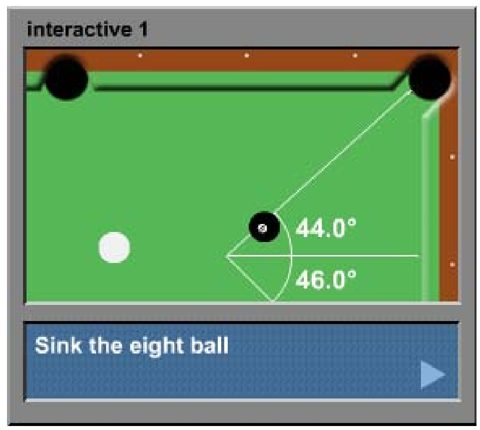
Step	Reason
3. $m\mathbf{v}_{i1} + m\mathbf{v}_{i2} = m\mathbf{v}_{f1} + m\mathbf{v}_{f2}$	conservation of momentum
4. $\mathbf{v}_{i1} = \mathbf{v}_{f1} + \mathbf{v}_{f2}$	simplify; $v_{i2} = 0$
5. $\mathbf{v}_{i1}\cdot\mathbf{v}_{i1} = (\mathbf{v}_{f1} + \mathbf{v}_{f2})\cdot(\mathbf{v}_{f1} + \mathbf{v}_{f2})$	dot product square of equation 4
6. $v_{i1}^2 = v_{f1}^2 + v_{f2}^2 + 2\mathbf{v}_{f1}\cdot\mathbf{v}_{f2}$	dot product of vector times itself

We now have two equations, one from conservation of energy and one from conservation of momentum. We subtract them and simplify.

Step	Reason
7. $2\mathbf{v}_{f1} \cdot \mathbf{v}_{f2} = 0$	subtract equation 2 from equation 6
8. $\mathbf{v}_{f1} \cdot \mathbf{v}_{f2} = 0$	divide by 2
9. $v_{f1}v_{f2} \cos(\theta + 50°) = 0$	dot product definition
10. $\cos(\theta + 50°) = 0$	divide by velocities
11. $\theta + 50° = \arccos(0) = 90°$	inverse cosine
12. $\theta = 40°$	evaluate

Notice that the angle made by the two balls after the collision is $90°$. This is true for any glancing elastic collision like this.

8.17 - Interactive problem: multi-dimensional collision



In the simulation at the right, you want to strike the stationary eight ball with the white cue ball so that it goes into the corner pocket. The cue ball will travel horizontally and strike the eight ball in an elastic collision, and then both balls will move in directions defined by the angles shown on the right. The friction on the table will slow the balls slightly.

You set what the speed of the cue ball will be at the instant it collides with the eight ball. You need to be careful about the resulting speed of the eight ball. It must travel fast enough to reach the pocket and sink, but if it is too fast, the cue ball will travel faster too, and you will scratch (sink the cue ball). The eight ball's speed is displayed on a gauge in the simulation. You want the speed of the eight ball immediately after the collision to be 1.00 m/s. The two balls have the same mass. Calculate what the speed of the cue ball should be at the moment of collision to sink the eight ball and not scratch the cue ball. Enter the speed to the nearest 0.01 m/s, and then press GO to verify your answer. Press RESET to try again.

You know that the total momentum is conserved, and the same is true for each component of the momentum. To solve this problem, you will need to apply conservation of momentum to the two directions, *x* and *y*, separately. Start by writing an equation for conservation of momentum in the *x*-dimension, which we will call the direction along which the cue ball initially travels.

$$mv_{ix,cue} + mv_{ix,8} = mv_{fx,cue} + mv_{fx,8}$$

Use trigonometry to express the velocity components in terms of each ball's speed and direction. The cue ball's initial velocity is solely in the *x*-direction and the eight ball's initial velocity is zero. In this problem, you can cancel the masses of the balls because they are identical.

$$v_{i,cue} = v_{f,cue} \cos\theta_{cue} + v_{f,8} \cos\theta_8$$

We will leave the rest up to you. Write a similar equation for conservation of momentum in the *y*-direction. Substitute the values given for the angles and for the eight ball's final speed into your system of equations, and solve for $v_{i,cue}$.

8.18 - Inelastic collisions

Inelastic collision: The collision results in a decrease in the system's total kinetic energy.

In an inelastic collision, momentum is conserved. But **kinetic** energy is not conserved. In inelastic collisions, kinetic energy transforms into other forms of energy. The kinetic energy after an inelastic collision is **less** than the kinetic energy before the collision. When one boxcar rolls and connects with another, as shown above, some of the kinetic energy of the moving car transforms into elastic potential energy, thermal energy and so forth. This means the kinetic energy of the system of the two boxcars decreases, making this an inelastic collision.

A *completely inelastic* collision is one in which two objects "stick together" after they collide, so they have a common final velocity. Since they may still be moving, completely inelastic does not mean there is zero kinetic energy after the collision. For instance, after the boxcars connect, the two "stick together" and move as one unit. In this case, the train combination continues to move after the collision, so they still have kinetic energy, although less than before the collision.

As with elastic collisions, we assume the collision occurs in an isolated system, with no net external forces present. You can think of elastic collisions and completely inelastic collisions as the extreme cases of collisions. Kinetic energy is not reduced at all in an elastic collision. In a completely inelastic collision, the total amount of kinetic energy after the collision is reduced as much as it can be, consistent with the conservation of momentum.

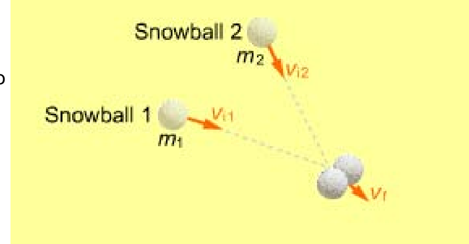
In Equation 1, you see an equation to calculate the final velocity of two objects, like the snowballs shown, after a completely inelastic collision. This equation is derived below. The derivation hinges on the two objects having a common velocity after the collision.

In this derivation, two objects, with masses m_1 and m_2, collide in a completely inelastic collision.

Step	Reason
1. $\mathbf{v}_{f1} = \mathbf{v}_{f2} = \mathbf{v}_f$	final velocities equal
2. $m_1\mathbf{v}_f + m_2\mathbf{v}_f = m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2}$	conservation of momentum
3. $\mathbf{v}_f(m_1 + m_2) = m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2}$	factor out v_f
4. $\mathbf{v}_f = \dfrac{m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2}}{m_1 + m_2}$	divide

Example 1 on the right applies this equation to a collision seen on many fall weekends: a football tackle.

equation 1



Completely inelastic collision

$$\mathbf{v}_f = \frac{m_1\mathbf{v}_{i1} + m_2\mathbf{v}_{i2}}{m_1 + m_2}$$

$\mathbf{v}_f$ = common final velocity
m_1, m_2 = masses
$\mathbf{v}_{i1}, \mathbf{v}_{i2}$ = initial velocities

example 1

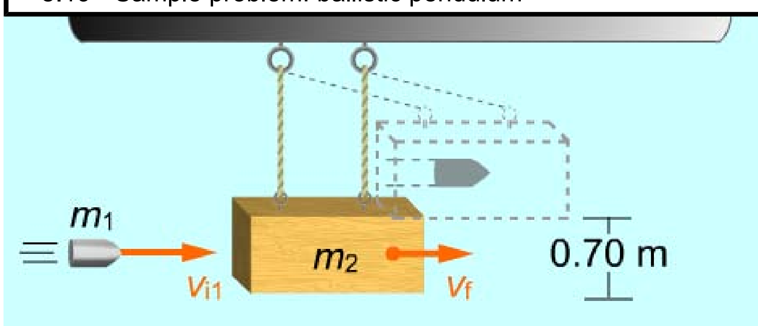
The players collide head-on at the velocities shown. What is their common velocity after this completely inelastic collision?

$$v_f = \frac{m_1 v_{i1} + m_2 v_{i2}}{m_1 + m_2}$$

$$v_f = \frac{(100\text{ kg})(6.4\text{ m/s}) + (92\text{ kg})(-5.6\text{ m/s})}{100 + 92\text{ kg}}$$

$\mathbf{v}_f$ = 0.65 m/s to the right

8.19 - Sample problem: ballistic pendulum



A bullet of mass 0.030 kg is fired into a wood block of mass 0.45 kg, and the block and bullet swing upward as shown. What was the initial velocity of the bullet?

A ballistic pendulum like the one you see above is used to measure the velocity of a fast moving projectile like a bullet. The bullet collides with the block, and the block and bullet together swing upward to a height that is noted.

We cannot simply use conservation of energy to equate the bullet's initial kinetic energy and the final potential energy of the system, because the collision is inelastic and some energy goes to deforming the wood. However, momentum is conserved in the collision. After the bullet embeds itself in the block, mechanical energy is conserved as the block swings upward. This means we first apply the principle of conservation of momentum and then the principle of conservation of energy to solve the problem.

Variables

Since this is a completely inelastic collision, the block and bullet have a common velocity just after the collision.

mass of bullet	$m_1 = 0.030$ kg
mass of block	$m_2 = 0.45$ kg
bullet's initial speed	v_{i1}
block's initial speed	$v_{i2} = 0$ m/s
speed of combined block/bullet just after collision	v_f
swing height of block/bullet	$h = 0.70$ m

What is the strategy?

1. Use conservation of momentum to write an equation relating the bullet's initial velocity to the common velocity of the block and bullet immediately after the collision.
2. Use conservation of energy to relate the potential energy of the block/bullet combination at its peak height to the kinetic energy of the combination right after the collision.
3. Combine these two equations to solve for the bullet's initial velocity.

Physics principles and equations

Since the initial and final velocities in the problem are confined to one dimension, we may write the equation for the conservation of momentum using signs to represent directions:

$$m_1v_{i1} + m_2v_{i2} = m_1v_{f1} + m_2v_{f2}$$

Definitions of kinetic energy and gravitational potential energy:

$$KE = ½mv^2$$

$$PE = mgh$$

Conservation of mechanical energy:

$$KE + PE = \text{constant}$$

Step-by-step solution

We start by using conservation of momentum to find an expression for the initial velocity of the bullet in terms of the common velocity of the block and bullet just after the collision.

Step	Reason
1. $m_1v_{i1}+m_2v_{i2}= m_1v_{f1}+m_2v_{f2}$	conservation of momentum
2. $m_1v_{i1} = m_1v_f + m_2v_f$	v_{i2} zero and $v_{f1} = v_{f2}$
3. $v_{i1} = \dfrac{m_1+m_2}{m_1}v_f$	solve for v_{i1}

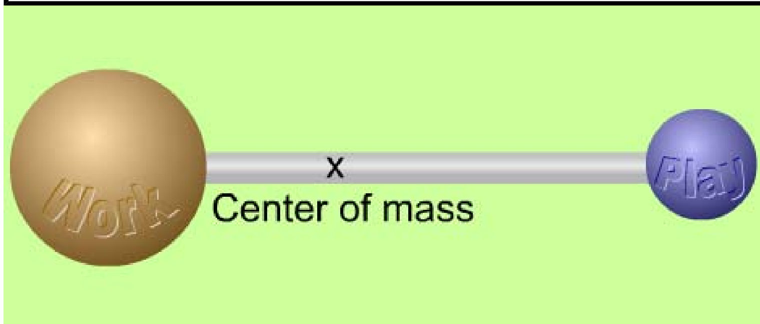
Now we find an expression for the common velocity of the block and bullet just after the collision, using the conservation of energy.

Step	Reason
4. $KE_{collision} = PE_{top}$	conservation of energy
5. $½(m_1+m_2)v_f^2 = (m_1+m_2)gh$	definitions of KE and PE
6. $½v_f^2 = gh$	simplify
7. $v_f = \sqrt{2gh}$	solve for v_f

Now we combine the results of the conservation of momentum and conservation of energy analyses to find the initial velocity of the bullet.

Step	Reason
8. $v_{i1} = \frac{m_1+m_2}{m_1}\sqrt{2gh}$	substitute equation 7 into equation 3
9. $v_{i1} = \frac{0.030+0.45\ \text{kg}}{0.030\ \text{kg}}\sqrt{2(9.80\ \text{m/s}^2)(0.70\ \text{m})}$	substitute values
10. $v_{i1} = 59$ m/s	evaluate

8.20 - Center of mass



concept 1

Center of mass
"Average" location of mass

Center of mass: Average location of mass. An object can be treated as though all its mass were located at this point.

The center of mass is useful when considering the motion of a complex object, or system of objects. You can simplify the analysis of motion of such an object, or system of objects, by determining its center of mass. An object can be treated as though all its mass is located at this point. For instance, you could consider the force of a weightlifter lifting the barbell pictured above as though she applied all the force at the center of mass of the bar, and determine the acceleration of the center of mass.

You may react: "But we have been doing this in many sections of this book," and yes, implicitly we have been. If we asked earlier how much force was required to accelerate this barbell, we assumed that the force was applied at the center of mass, rather than at one end of the barbell, which would cause it to rotate.

In this section, we focus on how to calculate the center of mass of a system of objects. Consider the barbell above. Its center of mass is on the rod that connects the two balls, nearer the ball labeled "Work," because that ball is more massive.

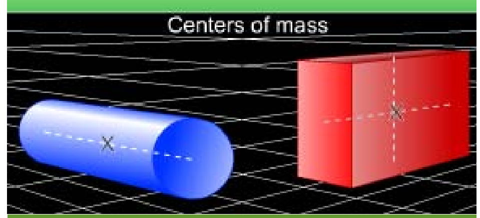
When an object is symmetrical and made of a uniform material, such as a solid sphere of steel, the center of mass is at its geometric center. So for a sphere, cube or other symmetrical shape made of a uniform material, you can use your sense of geometry and decide where the center of the object is. That point will be the center of the mass. (We can relax the condition of uniformity if an object is composed of different parts, but each one of them is symmetrical, like a golf ball made of different substances in spherically symmetrical layers.)

The center of mass does not have to lie inside the object. For example, the center of mass of a doughnut lies in the middle of its hole.

The equation to the right can be used to calculate the overall center of mass of a set of objects whose individual centers of mass lie along a line. To use the equation, place the center of mass of each object on the x axis. It helps to choose for the origin a point where one of the centers of mass is located, since this will simplify the calculation. Then, multiply the mass of each object times its center's x position and divide the sum of these products by the sum of the masses. The resulting value is the x position of the center of mass of the set of objects.

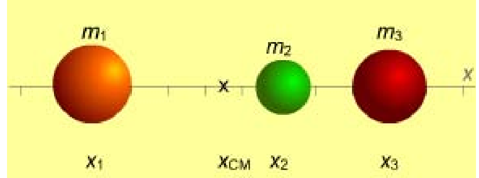
If the objects do not conveniently lie along a line, you can calculate the x and y positions of the center of mass by applying the equation in each dimension separately. The result is the x, y position of the system's center of mass.

concept 2



Center of mass
At geometric center of uniform, symmetric objects

equation 1

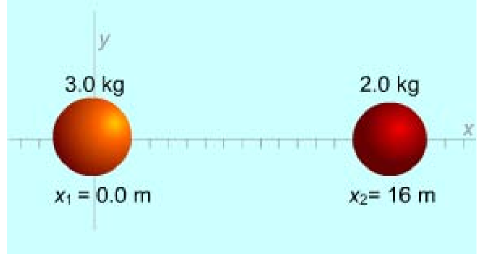


Center of mass

$$x_{CM} = \frac{m_1x_1+m_2x_2+\ldots+m_nx_n}{m_1+m_2+\ldots+m_n}$$

x_{CM} = x position of center of mass
m_i = mass of object i
x_i = x position of object i

example 1



What is the location of the center of mass?

$$x_{CM} = \frac{m_1x_1 + m_2x_2}{m_1 + m_2}$$

$$x_{CM} = \frac{(3.0kg)(0.0m) + (2.0kg)(16m)}{3.0kg + 2.0kg}$$

x_{CM} = 32/5.0

x_{CM} = 6.4 m

8.21 - Center of mass and motion

concept 1

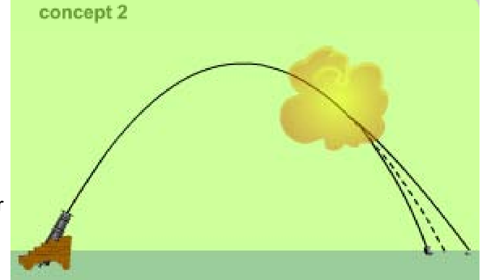
Center of mass and motion
Laws of mechanics apply to center of mass
Shifting center of mass creates "floating" illusion

Above, you see a ballet dancer performing a *grand jeté*, a “great leap.” When a ballet dancer performs this leap well, she seems to float through the air. In fact, if you track the dancer’s motion by noting the successive positions of her head, you can see that its path is nearly horizontal. She seems to be defying the law of gravity. This seeming physics impossibility is explained by considering the dancer’s center of mass. An object (or a system of objects) can be analyzed by considering the motion of its center of mass.

Look carefully at the locations of the dancer’s center of mass in the diagram. The center of mass follows the parabolic path of projectile motion.

To achieve the illusion of floating − moving horizontally − the dancer alters the location of her center of mass relative to her body as she performs the jump. As she reaches the peak of her leap, she raises her legs, which places her center of mass nearer to her head. This decreases the distance from the top of her head to her center of mass, so her head does not rise as high as it would otherwise. This allows her head to move in a straight line while her center of mass moves in the mandatory parabolic projectile arc.

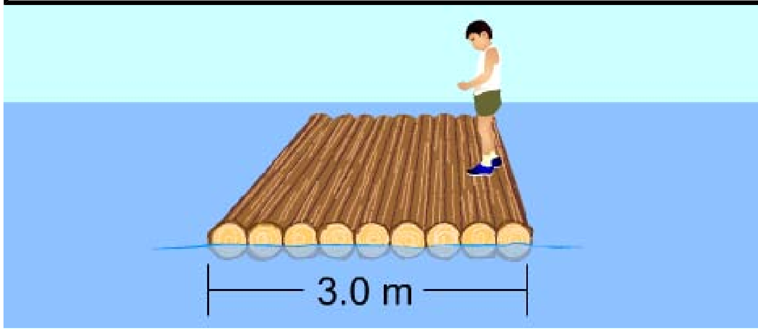
concept 2



Center of mass and motion
Center of mass follows projectile path

At the right, we use another example to make a similar point. A cannonball explodes in midair. Although the two resulting fragments move in different directions, the center of mass continues along the same trajectory the cannonball would have followed had it not exploded. The two fragments have different masses. The path of the center of mass is closer to the path of the more massive fragment, as you might expect.

8.22 - Sample problem: moving a raft

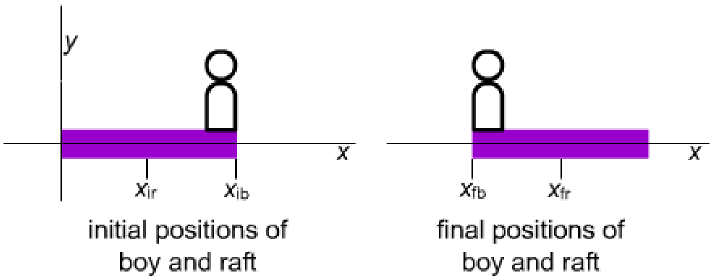


The 93 kg raft is initially stationary. A boy with mass 52 kg walks across it. How far does the raft move?

Above, you see a boy standing on one side of a raft made of wooden logs. The raft is stationary and no net external force acts on the boy/raft system (the water offers negligible resistance to its motion). If the boy walks to the other side of the raft, the raft will move. How far does it move?

Draw a diagram

To simplify matters, we place the left end of the raft at the origin for its initial position, so its center is at 1.5 m. After the boy walks, since we do not know the position of the raft, we do not draw the y axis in the illustration of the final positions.



Variables

	raft	boy
mass	$m_r = 93$ kg	$m_b = 52$ kg
initial position, center of mass	$x_{ir} = 1.5$ m	$x_{ib} = 3.0$ m
final position, center of mass	x_{fr}	$x_{fb} = x_{fr} - 1.5$ m

Note that we state the final position of the boy relative to the raft's center. He is 1.5 meters from it when he moves to the left side.

What is the strategy?

The key idea here is that **the center of mass of the boy/raft system does not move**. It does not move because it is initially at rest, and no net external force acts on the system. No net force means no acceleration, so the center of mass of the system must remain stationary.

1. Calculate the initial position of the center of mass of the system from the information provided above.
2. Write an equation for the center of mass of the system after the boy walks across the raft, enter values, and solve for the final position of the **raft's** center of mass. This uses the fact that the center of mass of the **system** does not move.
3. Calculate the raft's displacement from the initial and final positions of its center of mass.

Physics principles and equations

Equation to determine center of mass:

$$x_{CM} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}$$

The position of center of mass of the system remains constant since no external force acts on the system.

We also use the definition of displacement,

$$\Delta x = x_f - x_i$$

Step-by-step solution

First, we find the center of mass of the boy and raft before the boy walks across the raft. We place the left end of the raft at the origin, so its center is at x position 1.5 m and the boy is at 3.0 m.

Step	Reason
1. $x_{CM} = \dfrac{m_r x_{ir} + m_b x_{ib}}{m_r + m_b}$	definition of center of mass
2. $x_{CM} = \dfrac{(93)(1.5) + (52)(3.0)}{93 + 52}$	enter values
3. $x_{CM} = 2.0$ m	evaluate

The x position of the center of mass of the boy/raft system is 2.0 m. The position of the center of mass does not change after the boy walks across the raft, because there is no net external force on the system. We write an equation for calculating the center of mass with the boy on the other side of the raft, and use this equation to find the position of the raft. In step 5, we use the fact that the boy is 1.5 m to the left of the raft's center.

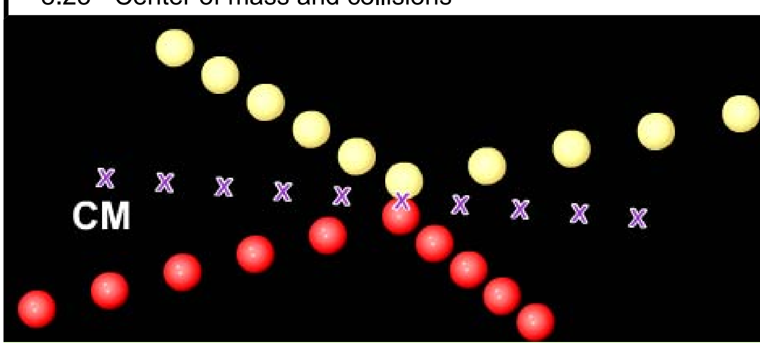
Step	Reason
4. $x_{CM} = \dfrac{m_r x_{fr} + m_b x_{fb}}{m_r + m_b}$	definition of center of mass
5. $x_{CM} = \dfrac{m_r x_{fr} + m_b (x_{fr} - 1.5)}{m_r + m_b}$	boy is left of center
6. $2.0 = \dfrac{93 x_{fr} + 52(x_{fr} - 1.5)}{93 + 52}$	enter values
7. $2.0 = \dfrac{145\, x_{fr} - 78}{145}$	simplify
8. $x_{fr} = 2.5$ m	solve for x_{fr}

Now that we know both the initial and final positions of the raft's center of mass, it is easy to find out how far it moved.

Step	Reason
9. $\Delta x_r = x_{fr} - x_{ir}$	definition of displacement
10. $\Delta x_r = 2.5\text{ m} - 1.5\text{ m}$	from 8 and initial condition
11. $\Delta x_r = 1.0$ m to the right	subtract

The raft moves one meter to the right when the boy walks across it to the left.

8.23 - Center of mass and collisions



concept 1

Collisions and center of mass

Center of mass has

· constant velocity

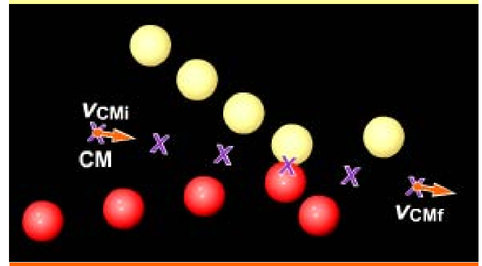
· constant momentum

Above, you see an illustration representing a time-lapse sequence of photographs of the elastic collision of two balls. Equal intervals of time elapse between each image of the balls. We have marked with x's the location of the center of mass at these intervals. The straight-line path and the equal distances between the x marks reflect a constant velocity. Since there is minimal net external force acting on **the system**, the center of mass does not accelerate but rather has essentially a constant velocity.

In general, when there is no **external** force on a system, the velocity of the center of mass is constant for any collision, elastic or not. In other words, since there is no net force on the system, there is no acceleration of the center of mass. We state this as the first equation on the right.

The momentum of the center of mass of this system is also constant. We state this as the second equation on the right. You can think of this as a consequence of the system's unchanging velocity and mass, or as an application of the general principle that momentum is conserved in a system when no net external force is present.

equation 1



When no external force is present

$\mathbf{v}_{CMi} = \mathbf{v}_{CMf}$

Velocity of CM does not change

$\mathbf{p}_{CMi} = \mathbf{p}_{CMf}$

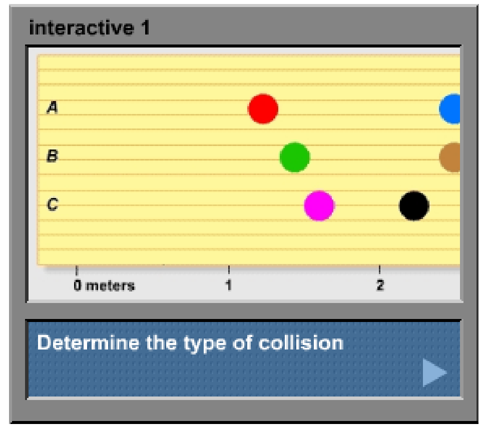
Momentum of CM does not change

8.24 - Interactive summary problem: types of collisions

On the right is a simulation featuring three collisions. Each collision is classified as one of the following: an elastic collision, a completely inelastic collision, an inelastic (but not completely inelastic) one, or an impossible collision that violates the laws of physics. The colliding disks all have the same mass, and there is no friction. Each disk on the left has an initial velocity of 1.00 m/s. The disks on the right have an initial velocity of −0.60 m/s.

Press GO to watch the collisions. Use the PAUSE button to stop the action after the collisions and record data, then make whatever calculations you need to classify each collision using the choices in the drop-down controls labeled "Collision type." Press RESET if you want to start the simulation from the beginning.

If you have difficulty with this, review the sections on elastic and inelastic collisions.



8.25 - Gotchas

One object has a mass of 1 kg and a speed of 2 m/s, and another object has a mass of 2 kg and a speed of 1 m/s. The two objects have identical momenta. Only if they are moving in the same direction. You **can** say they have equal magnitudes of momentum, but momentum is a vector, so direction matters. Consider what happens if they collide. The result will be different depending on whether they are moving in the same or opposite directions.

In inelastic collisions, momentum is not conserved. No. Kinetic energy decreases, but momentum **is** conserved.

Two objects are propelled by equal constant forces, and the second one is exposed to its force for three times as long. The second object's change in momentum must be greater than the first's. This is true. It experienced a greater impulse, and impulse equals the change in momentum.

8.26 - Summary

An object's momentum is the product of its mass and velocity. It is a vector quantity with units of kg·m/s.

Like energy, momentum is conserved in an isolated system. If no net external force acts on a system, its total momentum is constant.

A change in momentum is called impulse. It is a vector with the same units as momentum. Impulse can be calculated as the difference between the final and initial momenta, or as an average applied force times the duration of the force.

The conservation of momentum is useful in analyzing collisions between objects, since the total momentum of the objects involved must be the same before and after the collision. In an elastic collision, kinetic energy is also conserved. In an inelastic collision, the kinetic energy is reduced during the collision as some or all of it is converted into other forms of energy.

A collision is completely inelastic if the kinetic energy is reduced as much as possible, consistent with the conservation of momentum. The two objects "stick together" after the collision.

The center of mass of an object (or system of objects) is the average location of the object's (or system's) mass. For a uniform object, this is the object's geometric center. For more complicated objects and systems, center of mass equations must be applied.

Equations

$$\mathbf{p} = m\mathbf{v}$$

$$\Sigma\mathbf{F} = \frac{\Delta\mathbf{p}}{\Delta t}$$

$$\mathbf{J} = \overline{\mathbf{F}}\Delta t = \Delta\mathbf{p}$$

Conservation of momentum

$$\mathbf{p}_{i1} + \mathbf{p}_{i2} + \ldots + \mathbf{p}_{in} = \mathbf{p}_{f1} + \mathbf{p}_{f2} + \ldots + \mathbf{p}_{fn}$$

Center of mass

$$x_{CM} = \frac{m_1x_1 + m_2x_2 + \ldots + m_nx_n}{m_1 + m_2 + \ldots + m_n}$$

Moving objects behave as if all their mass were concentrated at their center of mass. For example, a hammer thrown into the air may rotate as it falls, but its center of mass will follow the parabolic path followed by any projectile.

Conceptual Problems

C.1 Two balls of equal mass move at the same speed in different directions. Are their momenta equal? Explain.

○ Yes ○ No

C.2 (a) If a particle that is moving has the same momentum and the same kinetic energy as another, must their masses and velocities be equal? If so, explain why, and if not, give a counterexample. (b) Describe the properties of two particles that have the same momentum, but different kinetic energies.

(a) ○ Yes ○ No

(b) Submit answer on paper

C.3 Two balls are moving in the same direction. Ball A has half the mass of ball B, and is moving at twice its speed. (a) Which ball has the greater momentum? (b) Which ball has greater kinetic energy?

(a)
- i. A's momentum is greater
- ii. B's momentum is greater
- iii. They have the same momentum
- iv. You cannot tell

(b)
- i. A's kinetic energy is greater
- ii. B's kinetic energy is greater
- iii. They have the same kinetic energy
- iv. You cannot tell

C.4 Two salamanders have the same mass. Their momenta (considered as signed quantities) are equal in magnitude but opposite in sign. Describe the relationship of their velocities.

C.5 A large truck collides with a small car. How does the magnitude of the impulse experienced by the truck compare to that experienced by the car?

- i. Truck impulse greater
- ii. Truck impulse equal
- iii. Truck impulse less

C.6 Object A is moving when it has a head-on collision with stationary object B. No external forces act on the objects. Which of the following situations are possible after the collision? Check all that are possible.

- ☐ A and B move in the same direction
- ☐ A and B move in opposite directions
- ☐ A moves and B is stationary
- ☐ A is stationary and B moves
- ☐ A and B are both stationary

C.7 A cannonball is on track to hit a distant target when, at the top of its flight, it unexpectedly explodes into two pieces that fly out horizontally. You find one piece of the cannonball, and you know the target location. What physics principle would you apply to find the other piece? Explain.

- i. Conservation of momentum
- ii. Impulse
- iii. Center of mass

C.8 An object can have a center of mass that does not lie within the object itself. Give examples of two such objects.

Section Problems

Section 0 - Introduction

0.1 Use the information given in the interactive problem in this section to answer the following questions. (a) Is it possible to have negative momentum? (b) Does the sum of the pucks' velocities remain constant before and after the collision? (c) Does the sum of the pucks' momenta remain constant?

(a) ○ Yes ○ No

(b) ○ Yes ○ No

(c) ○ Yes ○ No

Section 1 - Momentum

1.1 Belle is playing tennis. The mass of the ball is 0.0567 kg and its speed after she hits it is 22.8 m/s. What is the magnitude of the momentum of the ball?

__________ kg · m/s

1.2 A sport utility vehicle is travelling at a speed of 15.3 m/s. If its momentum has a magnitude of 32,800 kg·m/s, what is the SUV's mass?

__________ kg

1.3 A truck with a mass of 2200 kg is moving at a constant 13 m/s. A 920 kg car with the same momentum as the truck passes it at a constant speed. How fast is the car moving?

__________ m/s

1.4 At a circus animal training facility, a monkey rides a miniature motorscooter at a speed of 7.0 m/s. The monkey and scooter together have a mass of 29 kg. Meanwhile, a chimpanzee on roller skates with a total mass of 44 kg moves at a speed of 1.5 meters per second. The magnitude of the momentum of the monkey plus scooter is how many times the magnitude of the momentum of the chimpanzee plus skates?

____

1.5 A net force of 30 N is applied to a 10 kg object, which starts at rest. What is the magnitude of its momentum after 3.0 seconds?

__________ kg · m/s

Section 2 - Momentum and Newton's second law

2.1 A golden retriever is sitting in a park when it sees a squirrel. The dog starts running, exerting a constant horizontal force of 89 N against the ground for 3.2 seconds. What is the magnitude of the dog's change in momentum?

__________ kg · m/s

Section 3 - Impulse

3.1 A 1400 kg car traveling in the positive direction takes 10.5 seconds to slow from 25.0 meters per second to 12.0 meters per second. What is the average force on the car during this time?

__________ N

3.2 A cue stick applies an average force of 66 N to a stationary 0.17 kg cue ball for 0.0012 s. What is the magnitude of the impulse on the cue ball?

__________ kg · m/s

3.3 A baseball arrives at home plate at a speed of 43.3 m/s. The batter hits the ball along the same line straight back to the pitcher at 68.4 m/s. The baseball has a mass of 0.145 kg and the bat is in contact with the ball for 6.28×10^{-4} s. What is the magnitude of the average force on the ball from the bat?

__________ N

3.4 A steel ball with mass 0.347 kg falls onto a hard floor and bounces. Its speed just before hitting the floor is 23.6 m/s and its speed just after bouncing is 12.7 m/s. (a) What is the magnitude of the impulse of the ball? (b) If the ball is in contact with the floor for 0.0139 s, what is the magnitude of the average force exerted on the ball by the floor?

(a) __________ kg · m/s
(b) __________ N

3.5 The net force on a baby stroller increases at a constant rate from 0 N to 88 N, over a period of 0.075 s. The mass of the baby stroller is 7.4 kg. It starts stationary; what is its speed after the 0.075 s the force is applied?

__________ m/s

3.6 A government agency estimated that air bags have saved over 14,000 lives as of April 2004 in the United States. (They also stated that air bags have been confirmed as killing 242 people, and they stress that seat belts are estimated to save 11,000 lives a year.) Assume that a car crashes and has come to a stop when the air bag inflates, causing a 75.0 kg person moving forward at 15.0 m/s to stop moving in 0.0250 seconds. (a) What is the magnitude of the person's impulse? (b) What is the magnitude of the average force the airbag exerts on the person?

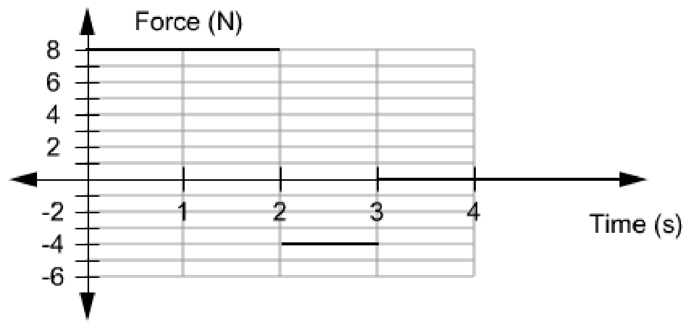
(a) __________ kg · m/s
(b) __________ N

3.7 Imagine you are a NASA engineer, and you are asked to design an airbag to protect the Mars Pathfinder from its impact with the Martian plain when it lands. Your system can allow an average force of up to 53,000 N on the spacecraft without damage. Your Pathfinder mock-up has a mass of 540 kg. If the spacecraft will strike the planet at 24 m/s, what is the minimum time for your airbag system to bring Pathfinder to rest so that the average force will not exceed 53,000 N?

__________ s

3.8 The graph shows the net force applied on a 0.15 kg object over a 3.0 s time interval. (a) What is the average force applied to the object over the 3.0 seconds? (b) What is the impulse? (c) What is its change in velocity?

(a) __________ N
(b) __________ kg · m/s
(c) __________ m/s



3.9 A ball is traveling horizontally over a volleyball net when a player "spikes" it, driving it straight down to the ground. The ball's mass is 0.22 kg, its speed before being hit is 6.4 m/s and its speed immediately after the spike is 21 m/s. What is the magnitude of the impulse from the spike?

__________ kg · m/s

3.10 For a movie scene, an 85.0 kg stunt double falls 12.0 m from a building onto a large inflated landing pad. After touching the landing pad surface, it takes her 0.468 s to come to a stop. What is the magnitude of the average net force on her as the landing pad stops her?

__________ N

3.11 A relative of yours belly flops from a height of 2.50 m (ouch!) and stops moving after descending 0.500 m underwater. Her mass is 62.5 kg. (a) What is her speed when she strikes the water? Ignore air resistance. (b) What is the magnitude of her impulse between when she hit the water, and when she stopped? (c) What was the magnitude of her acceleration in the pool? Assume that it is constant. (d) How long was she in the water before she stopped moving? (e) What was the magnitude of the average net force exerted on her after she hit the water until she stopped? (f) Do you think this hurt?

(a) __________ m/s
(b) __________ kg · m/s
(c) __________ m/s^2
(d) __________ s
(e) __________ N
(f) ○ Yes ○ No

3.12 Nitrogen gas molecules, which have mass 4.65×10^{-26} kg, are striking a vertical container wall at a horizontal velocity of positive 440 m/s. 5.00×10^{21} molecules strike the wall each second. Assume the collisions are perfectly elastic, so each particle rebounds off the wall in the opposite direction but at the same speed. (a) What is the change in momentum of each particle? (b) What is the average force of the particles on the wall?

(a) __________ kg · m/s
(b) __________ N

3.13 A pickup truck has a mass of 2400 kg when empty. It is driven onto a scale, and sand is poured in at the rate of 150 kg/s from a height of 3.0 m above the truck bed. At the instant when 440 kg of sand have already been added, what weight does the scale report?

__________ N

Section 6 - Conservation of momentum

6.1 A probe in deep space is infested with alien bugs and must be blown apart so that the icky creatures perish in the interstellar vacuum. The craft is at rest when its self-destruction device is detonated, and the craft explodes into two pieces. The first piece, with a mass of 6.00×10^{8} kg, flies away in a positive direction with a speed of 210 m/s. The second piece has a mass of 1.00×10^{8} kg and flies off in the opposite direction. What is the velocity of the second piece after the explosion?

__________ m/s

6.2 A 332 kg mako shark is moving in the positive direction at a constant velocity of 2.30 m/s along the bottom of a sea when it encounters a lost 19.5 kg scuba tank. Thinking the tank is a meal, it has lunch. Assuming momentum is conserved in the collision, what is the velocity of the shark immediately after it swallows the tank?

__________ m/s

6.3 A rifle fires a bullet of mass 0.0350 kg which leaves the barrel with a positive velocity of 304 m/s. The mass of the rifle and bullet is 3.31 kg. At what velocity does the rifle recoil?

__________ m/s

6.4 A cat stands on a skateboard that moves without friction along a level road at a constant velocity of 2.00 m/s. She is carrying a number of books. She wishes to stop, and does so by hurling a 1.20 kg book horizontally forward at a speed of 15.0 m/s with respect to the ground. (a) What is the total mass of the cat, the skateboard, and any remaining books? (b) What mass book must she now throw at 15.0 m/s with respect to the ground to move at −2.00 m/s?

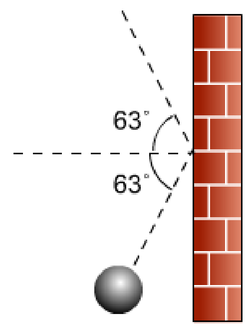
(a) __________ kg
(b) __________ kg

6.5 Stevie stands on a rolling platform designed for moving heavy objects. The platform has mass of 76 kg and is on a flat floor, supported by rolling wheels that can be considered to be frictionless. Stevie's mass is 43 kg. The platform and Stevie are stationary when she begins walking at a constant velocity of +1.2 m/s relative to the platform. (a) What is the platform's velocity relative to the floor? (b) What is Stevie's velocity relative to the floor?

(a) __________ m/s
(b) __________ m/s

Section 10 - Collisions

10.1 A steel ball of mass 0.76 kg strikes a brick wall in an elastic collision. Incoming, it strikes the wall moving 63° directly above a line normal (perpendicular) to the wall. It rebounds off the wall at an angle of 63° directly below the normal line. The ball's speed is 8.4 m/s immediately before and immediately after the collision, which lasts 0.18 s. What is the magnitude of the average force exerted by the ball on the wall?



__________ N

10.2 A quarterback is standing stationary waiting to make a pass when he is tackled from behind by a linebacker moving at 4.75 m/s. The linebacker holds onto the quarterback and they move together in the same direction as the linebacker was moving, at 2.60 m/s. If the linebacker's mass is 143 kg, what is the quarterback's mass?

__________ kg

Section 11 - Sample problem: elastic collision in one dimension

11.1 Two balls collide in a head-on elastic collision and rebound in opposite directions. One ball has velocity 1.2 m/s before the collision, and −2.3 m/s after. The other ball has a mass of 1.1 kg and a velocity of −4.2 m/s before the collision. (a) What is the mass of the first ball? (b) What is the velocity of the second ball after the collision?

(a) __________ kg
(b) __________ m/s

11.2 Two identical balls collide head-on in an elastic collision and rebound in opposite directions. The first ball has speed 2.3 m/s before the collision and 1.7 m/s after. What is the speed of the second ball (a) before and (b) after the collision?

(a) __________ m/s
(b) __________ m/s

11.3 A 4.3 kg block slides along a frictionless surface at a constant velocity of 6.2 m/s in the positive direction and collides head-on with a stationary block in an elastic collision. After the collision, the stationary block moves at 4.3 m/s. (a) What is the mass of the initially stationary block? (b) What is the velocity of the first block after the collision?

(a) __________ kg
(b) __________ m/s

11.4 Ball A, which has mass 23 kg, is moving horizontally left to right at +7.8 m/s when it overtakes the 67 kg ball B, which is moving left to right at +4.7 m/s, and they collide elastically. (a) What is ball A's velocity after the collision? (b) What is ball B's velocity?

(a) __________ m/s
(b) __________ m/s

Section 14 - Physics at play: elastic collisions and sports

14.1 Two steel balls are suspended on (massless) wires so that their centers align. One ball, with mass 2.30 kg, is pulled up and to the side so that it is 0.0110 m above its original position. Then it is released and strikes the other ball in an elastic collision. If the second ball has a mass of 3.10 kg, to what height does it rise above its original position?

_________ m

14.2 Ball 1 is moving at constant velocity when it strikes ball 2, which is initially stationary, in a head-on elastic collision. Prove that after the collision, the velocity of the ball 2 minus the velocity of ball 1 is the same as the initial velocity of ball 1.

Section 15 - Interactive problem: shuffleboard collisions

15.1 Use the information given in the interactive problem in this section to determine the initial velocity of the blue puck that will cause the red puck to stop at the scoring line. Test your answer using the simulation.

_________ m/s

Section 16 - Sample problem: elastic collision in two dimensions

16.1 Ball A collides with stationary ball B. Both balls have the same mass. After the collision, ball A has a speed of 2.7 m/s and moves at an angle of 17° from its original path of motion. Ball B has a speed of 2.2 m/s after the collision. (a) What is the angle (in degrees) between the directions of motion of the two balls after the collision? (b) What is the initial speed of ball A? (c) Is this an elastic collision?

(a) _________ °
(b) _________ m/s
(c) ○ Yes ○ No

16.2 On a frictionless surface, a 0.35 kg puck moves horizontally to the right (at an angle of 0°) and a speed of 2.3 m/s. It collides with a 0.23 kg puck that is stationary. After the collision, the puck that was initially moving has a speed of 2.0 m/s and is moving at an angle of −32°. What is the velocity of the other puck after the collision?

(_________ m/s, _________ °)

16.3 A 12.0 kg puck glides at 7.80 m/s over a frictionless surface at an angle of 0° and collides with a stationary 16.0 kg puck. After the collision, the 12.0 kg puck moves at an angle of +90°, at 3.20 m/s. What is the velocity of the 16.0 kg puck after the collision?"

(_________ m/s, _________ °)

16.4 In a game of marbles, a purple marble makes an off-center collision with a red marble of the same mass. Before the collision, the purple marble travels in the positive *x* direction and the red marble is at rest. After the collision, the purple marble's velocity of 0.0600 m/s makes a 35.0 degree angle above the positive *x* axis and the red marble's velocity makes an angle of negative 20.0 degrees with the axis. Find the initial speed of the purple marble.

_________ m/s

16.5 Two balls, both moving, hit each other in an elastic collision. Ball A has mass 4.00 kg and moves horizontally left to right at 3.00 m/s. Ball B has mass 5.00 kg and moves right and up at an angle of 40.0° from the horizontal. After the collision, ball A moves right and up at an angle of 20.0° from the horizontal and at a speed of 4.00 m/s. What are the direction and speed of ball B after the collision?

(_________ m/s, _________ °)

16.6 A ball of mass 11.0 kg moving at 3.00 m/s strikes a stationary ball of mass 12.0 kg in an off-center elastic collision. After the collision, the formerly stationary ball moves at an angle of 30.0° from the original path of the moving ball. (a) What is the speed of the 12.0 kg ball after the collision? (b) What is the speed of the 11.0 kg ball?

(a) _________ m/s
(b) _________ m/s

Section 17 - Interactive problem: multi-dimensional collision

17.1 Use the information given in the interactive problem in this section to calculate the initial velocity of the cue ball required to sink the eight ball. Test your answer using the simulation.

_________ m/s

Section 18 - Inelastic collisions

18.1 Ball A has mass 5.0 kg and is moving at −3.2 m/s when it strikes stationary ball B, which has mass 3.9 kg, in a head-on collision. If the collision is elastic, what is the velocity of (a) ball A, and (b) ball B after the collision? (c) If the collision is completely inelastic, what is the common velocity of balls A and B?

(a) __________ m/s
(b) __________ m/s
(c) __________ m/s

18.2 During a snowball fight, two snowballs travelling towards each other collide head-on. The first is moving east at a speed of 16.1 m/s and has a mass of 0.450 kg. The second is moving west at 13.5 m/s. When the snowballs collide, they stick together and travel west at 3.50 meters per second. What is the mass of the second snowball?

__________ kg

18.3 Three railroad cars, each with mass $2.3×10^4$ kg, are moving on the same track. One moves north at 18 m/s, another moves south at 12 m/s, and third car between these two moves south at 6 m/s. When the three cars collide, they couple together and move with a common velocity. What is their velocity after they couple?

__________ m/s

18.4 A 110 kg quarterback is running the ball downfield at 4.5 m/s in the positive direction when he is tackled head-on by a 150 kg linebacker moving at −3.8 m/s. Assume the collision is completely inelastic. (a) What is the velocity of the players just after the tackle? (b) What is the kinetic energy of the system consisting of both players before the collision? (c) What is the kinetic energy of the system consisting of both players after the collision?

(a) __________ m/s
(b) __________ J
(c) __________ J

18.5 Two clay balls of the same mass stick together in an completely inelastic collision. Before the collision, one travels at 5.6 m/s and the other at 7.8 m/s, and their paths of motion are perpendicular. If the mass of each ball is 0.21 kg, what is the magnitude of the momentum of the combined balls after the collision?

__________ kg · m/s

18.6 A large flat 3.5 kg boogie board is resting on the beach. Jessica, whose mass is 55 kg, runs at a constant horizontal velocity of 2.8 m/s. While running, she jumps on the board, and the two of them move together across the beach. (a) What is the speed of the board (with Jessica on it) just after she jumps on? (b) If the board and Jessica slide 7.7 m before coming to a stop, what is the coefficient of kinetic friction between the board and the beach?

(a) __________ m/s
(b) __________

18.7 A 6.0 kg ball A and a 5.0 kg ball B move directly toward each other in a head-on collision, then move in opposite directions away from the site of the collision. Ball A has velocity 4.1 m/s before the collision and −1.1 m/s after, and ball B has velocity −2.9 m/s before the collision. (a) What is ball B's velocity after the collision? (b) Is this an elastic collision?

(a) __________ m/s
(b) ○ Yes ○ No

Section 19 - Sample problem: ballistic pendulum

19.1 A dart gun suspended by strings hangs in equilibrium. The mass of the gun is 355 grams, not including a dart. The gun fires a 57.0 gram dart, causing it to swing backwards. The gun swings up to a height of 18.3 centimeters. What was the dart's speed in meters per second just after firing?

__________ m/s

19.2 A 0.0541 kg bullet is fired into a 3.25 kg block on a ballistic pendulum. The bullet goes straight through the block (without changing the mass of either object) and exits with a speed of 183 m/s. The block rises to a maximum height of 0.177 m. What was the initial speed of the bullet?

__________ m/s

19.3 A baseball (mass 0.145 kg) is at rest in a mesh net above the ground. A 0.023 kg marshmallow is thrown at the baseball from directly below, and attaches itself to the ball, without hitting the net, while traveling at a speed of 3.50 m/s. How high above its starting position does the baseball-marshmallow combination rise?

__________ m

19.4 A 1.30 kg book is resting on a horizontal surface. A large 0.120 kg spitball slides horizontally and sticks to the book. The book moves 0.320 m before coming to a rest. If the coefficient of kinetic friction between the book and the surface is 0.670, what was the speed of the spitball when it struck the book?

__________ m/s

Section 20 - Center of mass

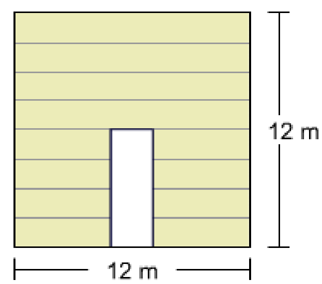
20.1 How far is the center of mass of the Earth-Moon system from the center of the Earth? The Earth's mass is 5.97×10^{24} kg, the Moon's mass is 7.4×10^{22} kg, and the distance between their centers is 3.8×10^{8} m.

__________ m

20.2 Four particles are postioned at the corners of a square that is 4.0 m on each side. One corner of the square is at the origin, one on the positive *x* axis, one in the first quadrant and one on the positive *y* axis. Starting at the origin, going clockwise, the particles have masses 2.3 kg, 1.4 kg, 3.7 kg, and 2.9 kg. What is the location of the center of mass of the system of particles?

(__________ , __________) m

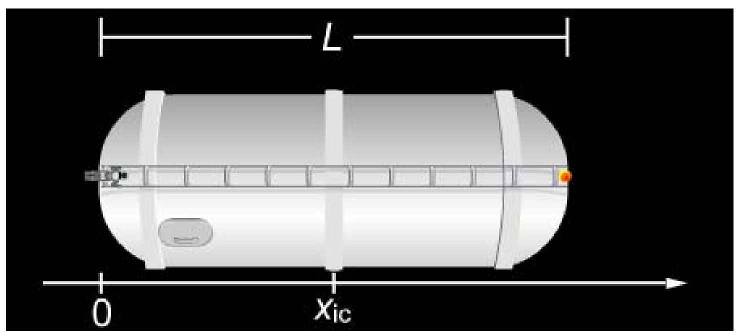
20.3 The uniform sheet of siding shown has a centrally-located doorway of width 2.0 m and height 6.0 m cut out of it. The sheet, with the doorway hole, has a mass of 264 kg. (a) What is the *x* coordinate of the center of mass? (b) What is the *y* coordinate of the center of mass? Assume the lower left hand corner of the siding is at (0, 0).



(a) __________ m
(b) __________ m

Section 22 - Sample problem: moving a raft

22.1 A future space traveller is in an escape capsule, far from any significant sources of gravity. To fix his emergency beacon, he climbs the entire length of a ladder, in the positive direction, that runs from one end of the capsule to the other. As he climbs, the capsule undergoes a displacement of −15.5 m. If the capsule has a mass of 5.00×10^{3} kg and the space traveller has a mass of 95.0 kg, what is the length of the capsule? Assume that the capsule's center of mass is halfway between its ends. Note that the art is not to scale.



__________ meters

Section 24 - Interactive summary problem: types of collisions

24.1 Use the information given in the interactive problem in this section to determine the collision type for (a) collision A, (b) collision B, and (c) collision C. Test your answer using the simulation.

(a)
i. Elastic
ii. Completely inelastic
iii. Inelastic
iv. Impossible

(b)
i. Elastic
ii. Completely inelastic
iii. Inelastic
iv. Impossible

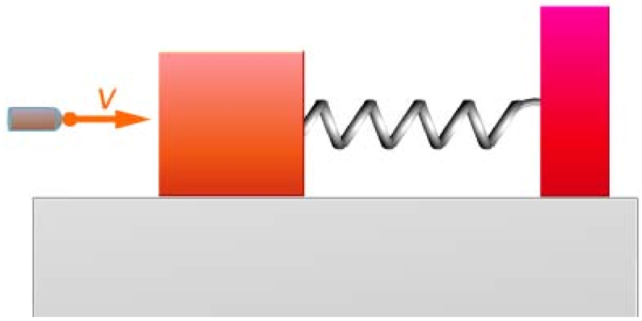
(c)
i. Elastic
ii. Completely inelastic
iii. Inelastic
iv. Impossible

Additional Problems

A.1 Two grains of space dust in a proto-planetary disk impact each other and stick together to form a larger chunk of space dust. Before the collision, grain one, with mass 4.60×10^{-9} kg, has a velocity of (3.00, 9.00) m/s and grain two, with mass 2.40×10^{-9} kg has a velocity of (0, 16) m/s. Find the final velocity of the resulting chunk of space dust in rectangular vector notation.

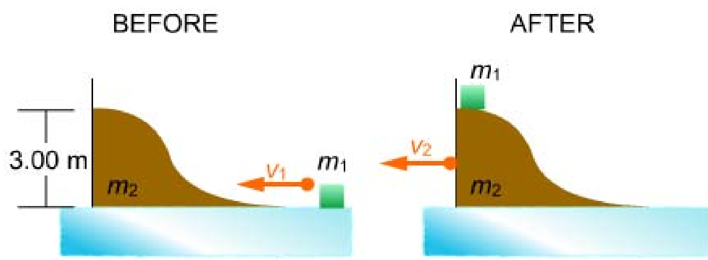
(__________ , __________) m/s

A.2 A 10.5 kg block, attached to the left end of a horizontal massless spring, sits on a frictionless table. The right end of the spring is attached to a vertical piece of wood that is firmly nailed to the table. A 0.0500 kg projectile is fired, from left to right, into the block at 85.5 m/s and stops inside it (this is a completely inelastic collision). The spring constant is k = 105 N/m. How many meters does the spring compress? The potential energy due to the compression of the spring can be calculated with the following formula: $PE = (1/2)kx^2$.



__________ m

A.3 A small block of mass m_1 = 0.500 kg slides to the left, with a velocity of v_1 = −8.40 m/s, towards a curved wedge of mass m_2 = 3.00 kg which sits on an icy, frictionless lake surface. When the block reaches the wedge, it slides up to a height 3.00 m above the ice. It comes to a stop relative to the wedge, and the two objects slide together to the left. (a) What is the velocity v_2 of the wedge-block system when they are sliding together? (b) Kinetic energy is not conserved during this process. What is the change in kinetic energy?



(a) __________ m/s

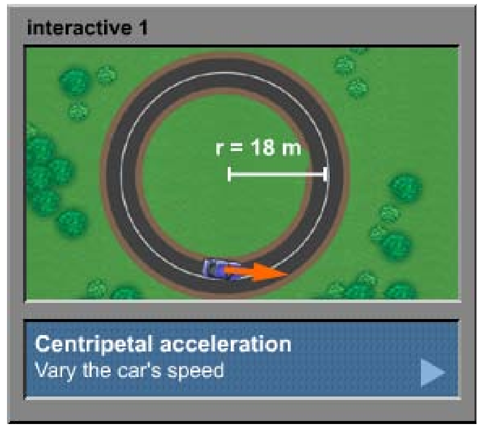
(b) __________ J

A.4 Ralphie's pet tortoise is stranded on a rock in the middle of a river. A rope swing is tied to a tree branch directly over the rock. Ralphie's plan is to swing from one river bank to the other, picking up his tortoise at the lowest point of the swing. Ralphie's mass is 40.0 kilograms and the tortoise's is 3.00 kg. If the opposite bank is 1.20 m above the tortoise, from what height will Ralphie have to start his swing if he is to reach the opposite bank? (This is an inelastic collision between Ralphie and his tortoise, so kinetic energy is not conserved during the collision.)

__________ meters

chapter 9 Uniform Circular Motion

9.0 - Introduction



A child riding on a carousel, you riding on a Ferris wheel: Both are examples of uniform circular motion. When the carousel or Ferris wheel reaches a constant rate of rotation, the rider moves in a circle at a constant **speed**. In physics, this is called uniform circular motion.

Developing an understanding of uniform circular motion requires you to recall the distinction between speed and velocity. Speed is the magnitude, or how fast an object moves, while velocity includes both magnitude and direction. For example, consider the car in the graphic on the right. Even as it moves around the curve at a constant speed, its velocity constantly changes as its direction changes. A change in velocity is called acceleration, and the acceleration of a car due to its change in direction as it moves around a curve is called centripetal acceleration.

Although the car moves at a constant speed as it moves around the curve, it is accelerating. This is a case where the everyday use of a word − acceleration − and its use in physics differ. A non-physicist would likely say: If a car moves around a curve at a constant speed, it is not accelerating. But a physicist would say: It most certainly is accelerating because its direction is changing. She could even point out, as we will discuss later, that a net external force is being applied on the car, so the car must be accelerating.

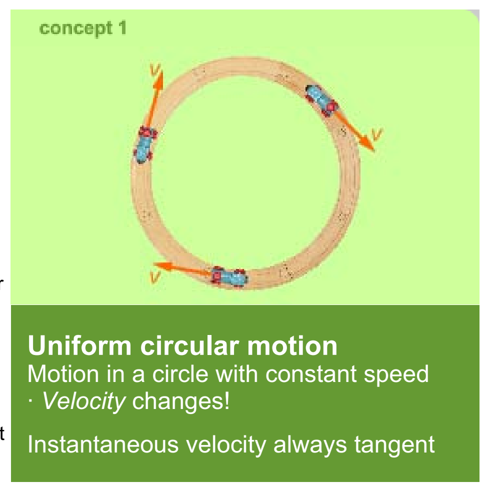
Uniform circular motion begins the study of rotational motion. As with linear motion, you begin with concepts such as velocity and acceleration and then move on to topics such as energy and momentum. As you progress, you will discover that much of what you have learned about these topics in earlier lessons will apply to circular motion.

In the simulation shown to the right, the car moves around the track at a constant speed. The red velocity vector represents the direction and magnitude of the car's instantaneous velocity.

The simulation has gauges for the x and y components of the car's velocity. Note how they change as the car travels around the track. These changes are reflected in the centripetal acceleration of the car. You can also have the car move at different constant speeds, and read the corresponding centripetal acceleration in the appropriate gauge. Is the centripetal acceleration of the car higher when it is moving faster? Note: If you go too fast, you can spin off the track. Happy motoring!

9.1 - Uniform circular motion

Uniform circular motion: Movement in a circle at a constant speed.



The toy train on the right moves on a circular track in uniform circular motion. The identical lengths of the velocity vectors in the diagram indicate a constant magnitude of velocity − a constant speed. When an object is moving in uniform circular motion, its speed is uniform (constant) and its path is circular.

The train does **not** have constant velocity; in fact, its velocity is constantly changing. Why? As you can also see in the diagram to the right, the direction of the velocity vector changes as the train moves around the track. A change in the direction of velocity means a change in velocity. The velocity vector is tangent to the circle at every instant because the train's displacement is tangent to the circle during every small interval of time.

Uniform circular motion is important in physics. For instance, a satellite in a circular orbit around the Earth moves in uniform circular motion.

9.2 - Period

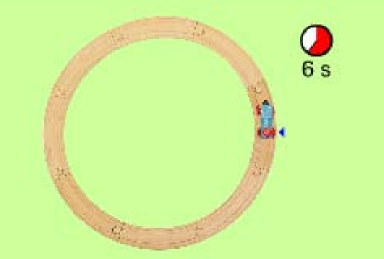
Period: The amount of time it takes for an object to return to the same position.

The concept of period is useful in analyzing motion that repeats itself. We use the example of the toy train shown in Concept 1 to illustrate a period. The train moves around a circular track at a constant rate, which is to say in uniform circular motion. It returns to the same position on the track after equal intervals of time. The period measures how long it takes the train to complete one revolution. In this example, it takes the train six seconds to make a complete lap around the track.

When an object like a train moves in uniform circular motion, that motion is often described in terms of the period. Many other types of motion can be discussed using the notion of a period, as well. For example, the Earth follows an elliptical path as it moves around the Sun, and its period is called a year. A metronome is designed to have a constant period that provides musicians with a source of rhythm.

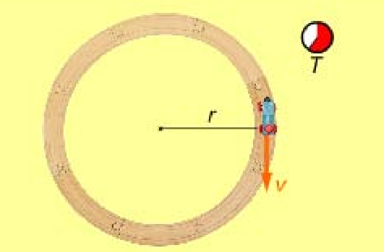
The equation on the right enables you to calculate the period of an object moving in uniform circular motion. The period is the circumference of the circle, $2\pi r$, divided by the object's speed. To put it more simply, it is distance divided by speed.

concept 1



Period
Time to complete one revolution

equation 1

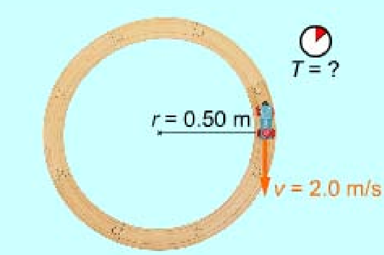


Period for uniform circular motion

$$T = \frac{2\pi r}{v}$$

T = period
r = radius
v = speed

example 1



What is the period of the train?

$$T = \frac{2\pi r}{v}$$

$$T = \frac{2\,(3.14)\,(0.50\ \text{m})}{(2.0\ \text{m/s})} = 1.6\ \text{s}$$

9.3 - Interactive checkpoint: a spinning CD

A CD player spins a CD at 205 revolutions per minute (rpm). The CD has a diameter of 12.0 cm. How fast is a point on the outer edge of the CD moving, in m/s?

Answer:

$v =$ ______ m/s

9.4 - Centripetal acceleration

Centripetal acceleration: The centrally directed acceleration of an object due to its circular motion.

An object moving in uniform circular motion constantly accelerates because its direction (and therefore its velocity) constantly changes. This type of acceleration is called centripetal acceleration. Any object moving along a circular path has centripetal acceleration.

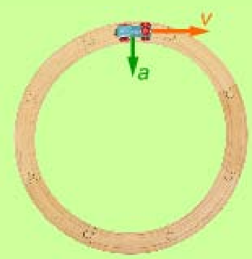
In Concept 1 at the right is a vector analysis of centripetal acceleration that uses a toy train as an example of an object moving along a circular path. As the drawing indicates, the train's velocity is tangent to the circle.

In uniform circular motion, the acceleration vector always points toward the center of the circle, perpendicular to the velocity vector. In other words, the object accelerates toward the center. This can be proven by considering the change in the velocity vector over a short period of time and using a geometric argument (an argument that is not shown here).

The equation for calculating centripetal acceleration is shown in Equation 1 on the right. The magnitude of centripetal acceleration equals the speed squared divided by the radius. Since both the speed and the radius are constant in uniform circular motion, the magnitude of the centripetal acceleration is also constant.

With uniform circular motion, the only acceleration is centripetal acceleration, but for circular motion in general, there may be both centripetal acceleration, which changes the object's direction, and acceleration in the direction of the object's motion (tangential acceleration), which changes its speed. If you ride on a Ferris wheel which is starting up, rotating faster and faster, you are experiencing both centripetal and tangential acceleration. For now, we focus on uniform circular motion and centripetal acceleration, leaving tangential acceleration as another topic.

concept 1



Centripetal acceleration

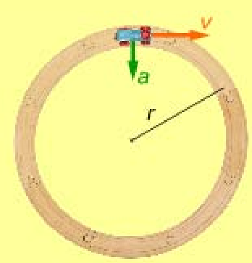
Acceleration due to change in direction in circular motion

In uniform circular motion, acceleration:

· Has constant magnitude

· Points toward center

equation 1



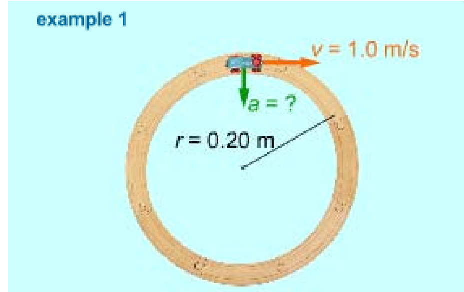
Centripetal acceleration

$$a_c = \frac{v^2}{r}$$

a_c = centripetal acceleration

v = speed

r = radius



What is the centripetal acceleration of the train?

$$a_c = \frac{v^2}{r}$$

$$a_c = \frac{1.0^2}{0.2} = 5.0\ \text{m/s}^2$$

Accelerates toward center

9.5 - Interactive problems: racing in circles

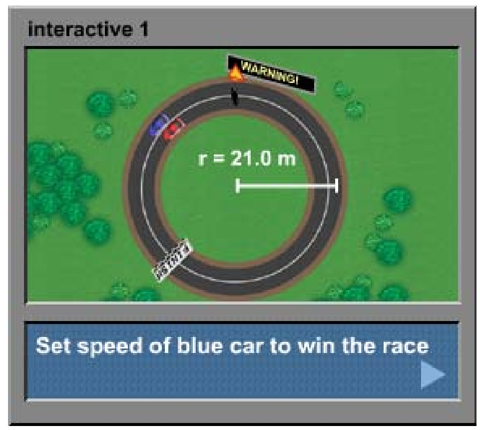
The two simulations in this section let you experience uniform circular motion and centripetal acceleration as you race your car against the computer's.

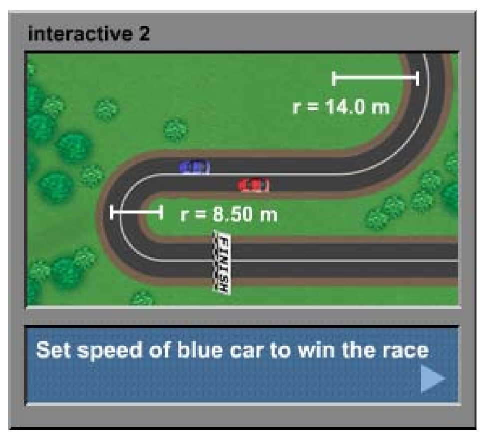
In the first simulation, you race a car around a circular track. Both your car and the computer's move around the loop at constant speeds. You control the speed of the blue car. Halfway around the track, you encounter an oil slick. If the centripetal acceleration of your car is greater than 3.92 m/s^2 at this point, it will leave the track and you will lose. The radius of the circle is 21.0 m.

To win the race, set the centripetal acceleration equal to 3.92 m/s^2 in the centripetal acceleration equation, solve for the velocity, and then round **down** the velocity to the nearest 0.1 m/s; this is a value that will keep your car on the track and beat the computer car. Enter this value using the controls in the simulation. Press GO to start the simulation and test your calculations.

In the second simulation, the track consists of two half-circle curves connected by a straight section. Your blue car runs the entire race at the speed that you set for it. You want to set this speed to just keep the car on the track. The first curve has a radius of 14.0 meters; the second, 8.50 meters. On either curve, if the centripetal acceleration of your car exceeds 9.95 m/s^2, its tires will lose traction on the curve, causing it to leave the track. If your car moves at the fastest speed possible without leaving the track, it will win. Again, calculate the speed of the blue car on each curve but using a centripetal acceleration value of 9.95 m/s^2, and round down to the nearest 0.1 m/s. Since the car will go the same speed on both curves, you need to decide which curve determines your maximum speed. Enter this value, then press GO.

If you have difficulty solving these interactive problems, review the equation relating centripetal acceleration, circle radius, and speed.





9.6 - Newton's second law and centripetal forces

If you hold onto the string of a yo-yo and twirl it in a circle overhead, as illustrated in Concept 1, you know you must hold the string firmly or the yo-yo will fly away from you. This is true even when the toy moves at a constant speed. A force must be applied to keep the yo-yo moving in a circle.

A force is required because the yo-yo is accelerating. Its change in direction means its velocity is changing. Using Newton's second law, $F = ma$, we can calculate the amount of force as the product of the object's mass and its centripetal acceleration. That equation is shown in Equation 1. It applies to any object moving in uniform circular motion. The force, called a *centripetal force*, points in the same direction as the acceleration, toward the center of the circle.

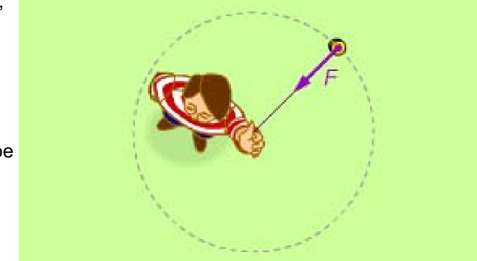
The term "centripetal" describes any force that causes circular motion. A centripetal force is not a new type of force. It can be the force of tension exerted by a string, as in the yo-yo example, or it can be the force of friction, such as when a car goes around a curve on a level road.

It can also be a normal force; for example, the walls of a clothes dryer supply a normal force that keeps the clothes moving in a circle, while the holes in those walls allow water to “spin out” of the fabric. Or, as in the case of the motorcycle rider in Example 1, the centripetal force can be a combination of forces, such as the normal force from the wall and the force of friction.

Sometimes the source of a centripetal force is easily seen, as with a string or the walls of a dryer. Sometimes that force is invisible: The force of gravity cannot be directly seen, but it keeps the Earth in its orbit around the Sun. The centripetal force can also be quite subtle, such as when an airplane tilts or banks; the air passing over the plane’s angled wings creates a force inward. In each of these examples, a force causes the object to accelerate toward the center of its circular path.

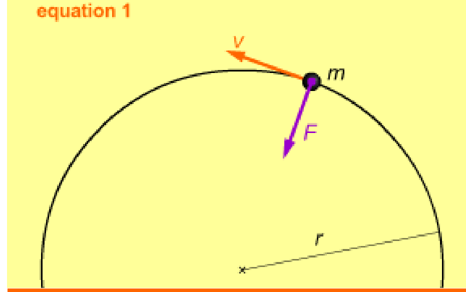
Identifying the force or forces that create the centripetal acceleration is a key step in solving many problems involving circular motion.

concept 1



Forces and centripetal acceleration
Force causes circular motion
Directed toward center
Any force can be centripetal

equation 1

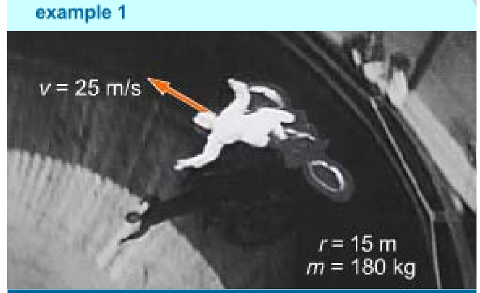


Forces and centripetal acceleration

$$F = m\frac{v^2}{r}$$

F = net force
m = mass
v = speed
r = radius

example 1

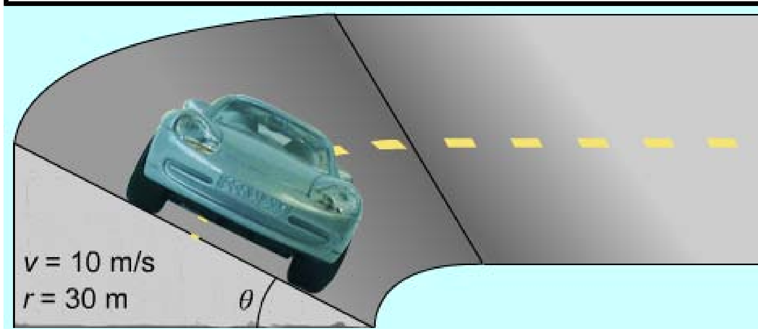


A daredevil bike rider goes around a circular track. The bike and rider together have the mass shown. What is the centripetal force on them?

$$F = m\frac{v^2}{r}$$

$F = (180\text{ kg})(25\text{ m/s})^2/15\text{ m}$
$F = (180)(625)/15$
$F = 7500\text{ N}$

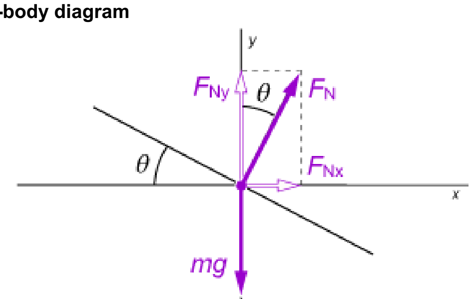
9.7 - Sample problem: banked curves



The car goes around a banked circular track. What bank angle θ will let it negotiate the curve without any friction on the tires?

Curves on roads are often banked, which is to say that they are tilted at an angle instead of flat. The question above asks for the angle at which the car can go around a circular curve without requiring friction. The normal force of the road supplies the force necessary to create the centripetal acceleration. This may at first seem implausible, but the analysis in this section shows how it can be done.

Draw a free-body diagram



We draw a free-body diagram of the forces on the car, assuming there is no friction. The banking of the curve is designed to keep the car on course. The normal force is perpendicular to the road surface, so it is at the same angle θ from the vertical as the bank angle.

Variables

speed	v = 10 m/s	
radius	r = 30 m	
bank angle	θ	
	x component	y component
normal force	$F_{Nx} = F_N \sin\theta$	$F_{Ny} = F_N \cos\theta$
weight	0	$mg \sin 270° = -mg$

What is the strategy?

1. The horizontal acceleration of the car is centripetal acceleration. Apply Newton's second law, setting the horizontal force equal to the car's mass times its centripetal acceleration.
2. Since the car does not accelerate in the vertical direction, the sum of the vertical forces (the vertical component of the normal force and the weight of the car) is zero. State this as an equation.
3. Combine these equations and simplify.

Physics principles and equations

Newton's second law

$$\Sigma F = ma$$

The equation for centripetal acceleration

$$a_c = v^2/r$$

Step-by-step solution

The first series of steps uses the diagram above.

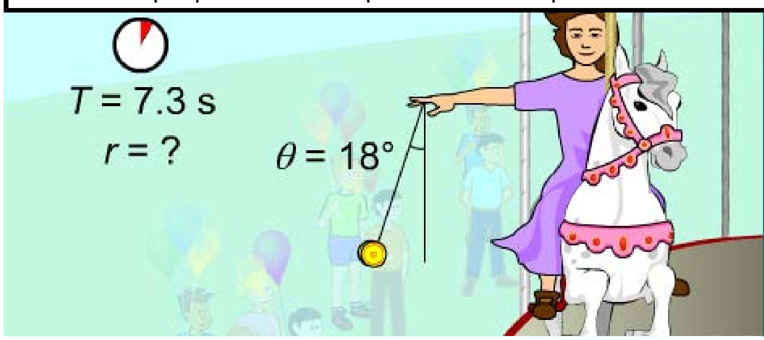
Step	Reason
1. $\Sigma F_x = F_N \sin\theta = mv^2/r$	The horizontal component of the normal force is the centripetal force
2. $\Sigma F_y = F_N \cos\theta + (-mg) = 0$	Newton's second law; no acceleration in vertical dimension
3. $F_N \cos\theta = mg$	rewrite
4. $\frac{F_N \sin\theta}{F_N \cos\theta} = \frac{mv^2}{rmg}$	divide equation 1 by equation 3
5. $\frac{\sin\theta}{\cos\theta} = \frac{v^2}{rg}$	simplify
6. $\tan\theta = \frac{v^2}{rg}$	trigonometric identity

This gives us an equation we can use to solve the problem.

Step	Reason
7. $\theta = \arctan(v^2/rg)$	solve equation 6 for θ
8. $\theta = \arctan\left(\frac{(10\text{ m/s})^2}{(30\text{ m})(9.80\text{ m/s}^2)}\right)$	enter values
9. $\theta = \arctan(0.34)$	arithmetic
10. $\theta = 19°$	value of arctan

The perhaps surprising and definitely fortunate result is that the bank angle does not depend on the mass of the vehicle. For a given speed and radius, the same angle will work for a tricycle or a truck. This explains why road curves have speed limit signs that announce the maximum safe speed for a vehicle without having to specify its mass.

9.8 - Sample problem: centripetal force on a pendulum



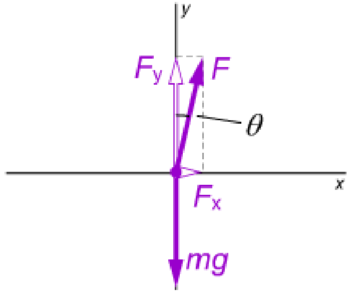
The yo-yo swings outward at the angle shown as the carousel rotates. What is the radius of the carousel from its center to the yo-yo?

Your friend decides to dangle a yo-yo from its string while riding a carousel that rotates at a constant speed. The yo-yo swings outward. Knowing the angle at which the yo-yo hangs, and the period of the carousel's rotation, you can find the radius of the carousel, as the problem asks.

To solve this problem, you need to consider the source of the centripetal force on the yo-yo. Since the yo-yo hangs at an angle, there is a horizontal component of the string tension. The horizontal component of tension provides the force for centripetal acceleration of the yo-yo.

Draw a free-body diagram

A free-body diagram of the yo-yo can be used to analyze the forces. We use F for the tension force.



Variables

period	$T = 7.3$ s	
angle of yo-yo	$\theta = 18°$	
radius	r	
speed of yo-yo	v	
	x component	y component
tension	$F_x = F \sin \theta$	$F_y = F \cos \theta$
weight of yo-yo	0	$mg \sin 270° = -mg$

What is the strategy?

1. Using Newton's second law, write an equation stating that the horizontal component of the tension creates the centripetal acceleration of the yo-yo.
2. The yo-yo does not accelerate vertically. State this as another equation using Newton's second law.
3. Combine and simplify the equations.
4. Use the relationship of period and speed to eliminate the speed of the yo-yo from the resulting equation.
5. Simplify and solve the equation.

Physics principles and equations

Newton's second law

$$\Sigma F = ma$$

The equation for centripetal acceleration

$$a_c = v^2/r$$

An equation for the period of an object in uniform circular motion

$$T = 2\pi r/v$$

Step-by-step solution

We first use Newton's second law to find an equation for the string angle involving the radius and the yo-yo's speed.

Step	Reason
1. $\Sigma F_x = F \sin \theta = mv^2/r$	The x component of the tension is the centripetal force
2. $\Sigma F_y = F \cos \theta + (-mg) = 0$	Newton's second law; no acceleration in vertical dimension
3. $F \cos \theta = mg$	rewrite
4. $\dfrac{F\sin\theta}{F\cos\theta} = \dfrac{mv^2}{rmg}$	divide equation 1 by equation 3
5. $\dfrac{\sin\theta}{\cos\theta} = \dfrac{v^2}{rg}$	simplify
6. $\tan\theta = \dfrac{v^2}{rg}$	trigonometric identity

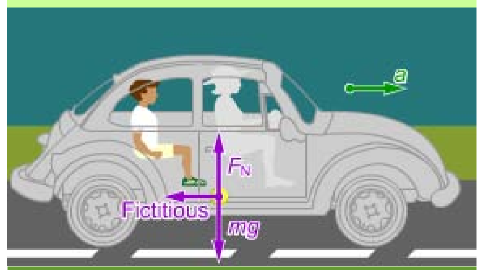
We created an equation that has two terms we do not know: the radius and the speed of the yo-yo. We use the relationship of period and speed to reduce the equation to one unknown term, the radius, and solve.

Step	Reason
7. $T = 2\pi r/v$	period of object in uniform circular motion
8. $v = 2\pi r/T$	solve for v
9. $\tan\theta = \dfrac{(2\pi r)^2}{T^2 rg}$	substitute for v in equation 6
10. $r = \dfrac{T^2 g \tan\theta}{4\pi^2}$	solve for r
11. $r = \dfrac{(7.3\ \text{s})^2(9.80\ \text{m/s}^2)\tan 18°}{4(3.14)^2}$	enter values
12. $r = 4.3$ m	evaluate

9.9 - Accelerating reference frames and fictitious forces

Fictitious force: A perception of force caused by the acceleration of a reference frame.

concept 1



Accelerating reference frame
Objects appear to violate Newton's laws
Fictitious forces seem to reconcile violation

Imagine that you are riding in the back seat of a car, heading home after a day of tennis. Suddenly, you see a tennis ball on the floor roll toward the back of the car while you feel yourself moving back and the seat pressing harder against you.

From experience, you know what has happened: The driver has pressed her foot down on the gas pedal, increasing the car's velocity. She caused the car to accelerate, which in turn caused the ball to roll and created the pushing sensation you experienced.

If you were to consider this solely from the reference frame of the car, you would be hard-pressed to explain the origin of the force that accelerated the ball. In this reference frame, no force can be identified to explain why the ball accelerated toward you, because no force is being applied on it. Of course, you could look outside, but consider perhaps that it is a dark night, or you are instead in a spacecraft, where reference points are not as readily viewed. You can explain what is occurring by using a different reference frame, but we are focusing on what you perceive while inside the car, with its interior defining your reference frame.

concept 2

Non-accelerating reference frame
Car's acceleration explains observed motion

A force diagram using the car as the reference is shown in Concept 1. The normal force of the floor "up" balances the weight of the ball "down" so there is no net force in the vertical direction, and the ball is not accelerating up or down. This all fits with your understanding of physics. However, in this frame of reference the ball is accelerating backward with no apparent force on it. This defies Newton's second law: There is acceleration but no net force. So, you may decide to look for an explanation.

To explain this, you can do one of two things. First, you can use what physicists call a *fictitious force* (sometimes called a *pseudoforce*). You can imagine that some force pushes you and the ball backward. This new force is called a fictitious force because there is no true force you can point to (such as gravity or a normal force) to explain it. Repeat: This is a fictitious force, and by fictitious, we mean one that truly does not exist. If you leave this chapter thinking there is some force pushing the ball backward, we have failed! The ball moves backward because the reference frame (the car) is accelerating, and Newtonian mechanics do not hold in an accelerating reference frame.

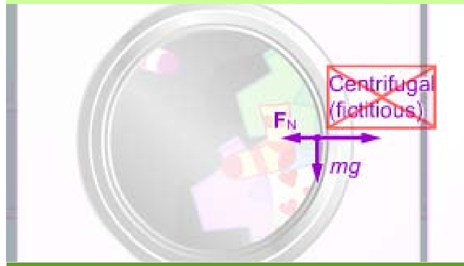
To explain the ball's acceleration, you could also change the reference frame to one that is **not** accelerating, such as the one used by an observer standing on the ground, as shown in Concept 2. (This reference frame accelerates slightly due to the Earth's motion, but we ignore this here.) From this person's perspective, the car is accelerating forward, as is the passenger inside. The normal force of the seat causes the person inside to accelerate. The ball is not accelerating, which is why it moves backward relative to the car. The ball moves at a constant velocity, not accelerating, because there is no net force acting on it.

Circular motion provides several familiar situations involving fictitious forces. Consider the clothes in a dryer as shown in Concept 3. As the dryer spins, the clothes press against its walls. Does some outward force, a *centrifugal force*, cause this? No. In fact, the wall of the dryer pushes against the clothes inward, supplying the inward centripetal force. (A force of friction not shown in the diagram also results from this normal force.) The dryer supplies both the mechanism to put the clothes in motion and the force to keep them from exiting.

Clothes presumably do not consider the acceleration of their reference frames, but you factor this into your behavior at times. For example, when you are in a car that is traveling around a curve in the road, you may feel as if there is a force pulling you outward. And, if the car is moving fast enough or the curve is particularly tight, you might push against the car with your foot or hand to stabilize yourself. Doing so increases the normal force of the car on your body, causing your body to remain in a circular path.

Despite the efforts of physicists, the concept of centrifugal force is hard to eradicate. Perhaps explaining accelerating reference frames is just too hard. For instance, a U.S. government press release about the safety of a carnival ride states: "The ride is a whirling cylinder which uses centrifugal force to hold the riders to their seats as the seats rise…."

concept 3



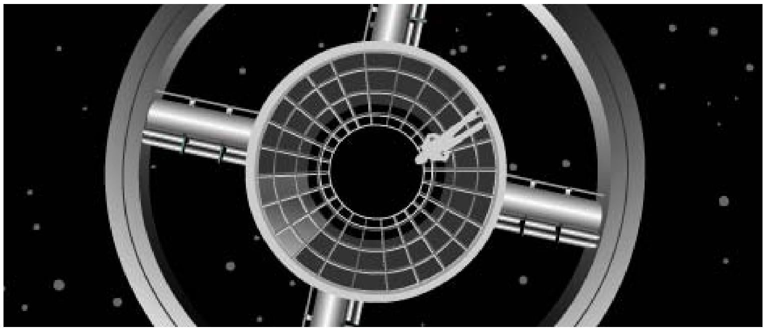
Accelerating reference frame, circular
Circular motion linked to fictitious forces
· Outward centrifugal "force" (fictitious)

Circular motion requires centripetal force

9.10 - Artificial gravity

When a spacecraft is far away from the Earth and any other massive body, the force of gravity is near zero. As a result of this lack of gravitational force, the astronauts and their equipment float in space.

Although perhaps amusing to watch and experience, floating presents an unusual challenge because humans are accustomed to working in environments with enough gravity to keep them, and their equipment, anchored to the floor. (Note: This same feeling of weightlessness occurs in a spacecraft orbiting the Earth, but the cause of the apparent weightlessness in this case arises from the free-fall motion of the craft and the astronaut, not from a lack of gravitational force.)



A rotating space station provides an illusion of gravity.

A number of science fiction books and films have featured spacecraft that rotate very slowly as they travel through the universe. Arthur C. Clarke's science fiction novel *Rendezvous with Rama* is set in a massive rotating spacecraft, as is part of the movie *2001*, which is also based on his work. This rotation supplies *artificial gravity* − the illusion of gravity − to the astronauts and their equipment. Artificial gravity has effects similar to true gravity, and as a result can mislead the people riding in such machines to believe they are experiencing true gravitational force.

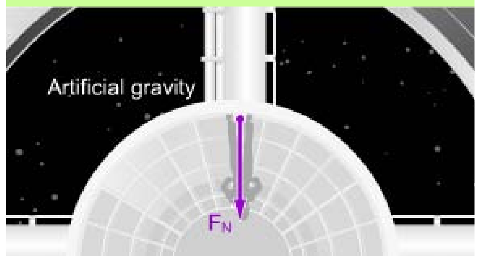
Why does the rotation of a spacecraft produce the **sensation** of gravity? Consider what happens when an airplane takes off from a runway: You feel a force pulling you back into your seat, as if the force of gravity were increasing. The force of gravity has not been significantly altered (in fact, it decreases a bit as you gain elevation). However, while the airplane accelerates upward, you feel a greater normal force pushing up from your seat, and you may interpret this subconsciously as increased gravity.

A roughly analogous situation occurs on a rotating spaceship. The astronauts are rotating in uniform circular motion. The outside wall of the station (the floor, from the astronauts' perspective) provides the centrally directed normal force that is the centripetal force. This force keeps the astronauts moving in a circle. From the astronauts' perspective, this force is upwards, and they relate it to the upward normal force of the ground they feel when standing on the Earth. On Earth, the normal force is equal but opposite to the force of gravity. Because they typically associate the normal force with gravity, the astronauts may erroneously perceive this force from the spacecraft floor as being caused by some form of artificial or simulated gravity.

Artificial gravity is a pseudo, or fictitious, force. The astronauts assume it exists because of the normal force. The perception of this fictitious force is a function of the acceleration of the astronauts as they move in a circle. It would disappear if the spacecraft stopped rotating.

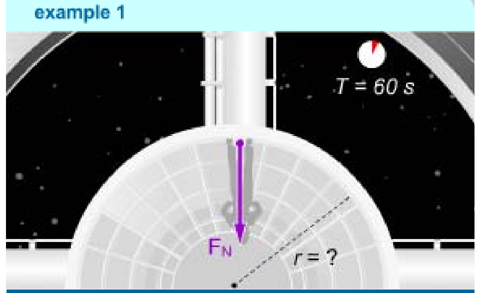
Although discussed as the realm of science fiction, real-world carnival rides (like the "Gravitron") use this effect. Riders are placed next to the wall of a cylinder. The cylinder then is rotated at a high speed and the floor (or seats) below the riders is lowered. The walls of the cylinder supply a normal force and the force of friction keeps the riders from

concept 1



Artificial gravity
Space station rotates
Floor of craft provides centripetal force
Person (incorrectly) assumes normal force counters force of gravity

example 1



To simulate Earth's gravity, what should the radius of the space

slipping down.

station be?

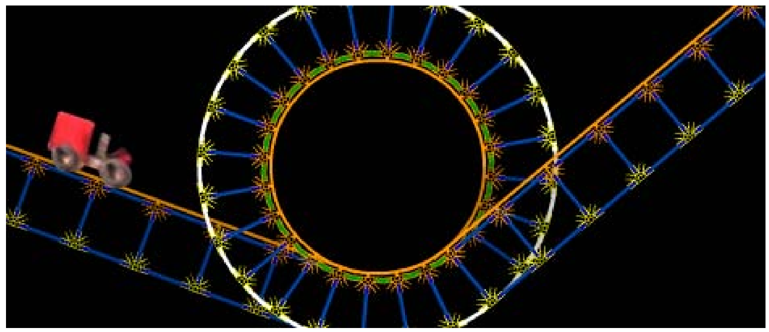
$$\frac{v^2}{r} = g \text{ and } v = \frac{2\pi r}{T}$$

$$\frac{4\pi^2 r^2}{T^2 r} = g$$

$$r = \frac{gT^2}{4\pi^2}$$

$$r = \frac{(9.80\ \text{m/s}^2)(60\ \text{s})^2}{(4)(3.14^2)} = 890\ \text{m}$$

9.11 - Loop-the-loop



A toy car loops the loop.

Above, our toy is successfully completing a circle around the loop, in the process "defying" gravity. This loop-the-loop scenario provides another opportunity to apply the concept of centripetal force.

We ignore friction and consider only the gravitational force and the normal force on the toy from the track. The normal force always points toward the center of the loop, while the gravitational force always points down. Consequently, the net force on the toy and its speed change as it goes around the loop. For example, the normal force is horizontal when the toy is at the 3 o'clock position and vertical at the top of the loop.

We apply Newton's second law to the toy at the top of the loop, where the gravitational and normal forces both point down toward the center of the loop. At that point, the sum of the normal and gravitational forces equals the toy's mass times its centripetal acceleration.

At the top of the loop, we will assume the toy is moving at the minimum speed required to keep from falling off the track. At this minimum speed, the normal force from the track is zero. This means the only force acting on the toy at this point is gravity; it provides all the centripetal force.

To derive the equation for the minimum speed required to negotiate the top of the loop, we set the weight equal to the mass times the centripetal acceleration. By re-arranging this equation, we can solve for the speed, as we show in Equation 1 on the right. Note that the toy's mass is not a factor in this equation. To apply this equation precisely, the radius is that of the circle of motion of the car's center of mass, not the radius of the track itself.

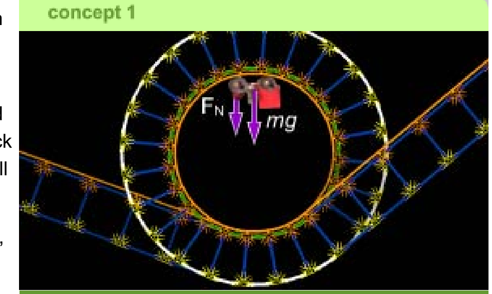
Since in most cases the center of mass of the car is close to the loop's track, the radius of the loop is often used instead. For the toy on the right, this is a significant approximation, but for typical roller coasters, it is much more reasonable.

If the toy moves faster at the top of the loop than the speed calculated in the equation, the centripetal acceleration will be greater. The normal force of the track will supply the additional force required for the increased acceleration. If the toy is moving more slowly than the speed calculated on the right, it will lose contact with the track.

As mentioned, the equation on the right shows how to calculate the minimum speed at the top of the loop required for a given radius of the object's motion. The greater the radius, the greater the speed required. To give you an idea of the speed required for navigating a loop built on a larger scale, the minimum speed for a radius of 10 meters is 9.9 m/s (about 36 km/h).

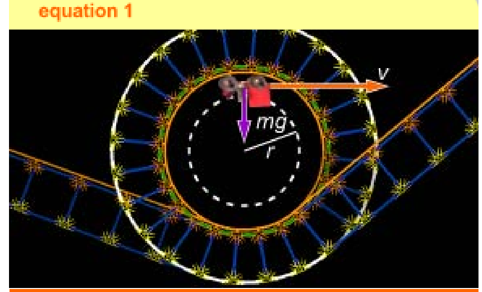
Although roller coasters are lightly regulated, efforts are underway to legislate a maximum *G force*. This force represents the rider's perceived weight as a multiple of his normal weight. When experiencing a force of 2 G, the rider feels twice as heavy as usual. The perceived weight is a combination of the rider's actual weight and the forces creating the acceleration of the roller coaster. Suggested maximum G forces range from 2.5 G to 3 G. Here is an example to give you an idea of the accelerations involved. For ease of comparison, we assume the same speed at the top and bottom of the loop, though this would not be true for a real coaster. A person moving in a vertical loop with a radius of 10 meters, at a constant speed of 16 m/s (58 km/h), experiences a 1.6 G force at the top of the loop but a sensational 3.6 G force at the bottom.

concept 1



At top of loop:
Weight plus any normal force are centripetal force(s)
At minimum speed, normal force is zero

equation 1



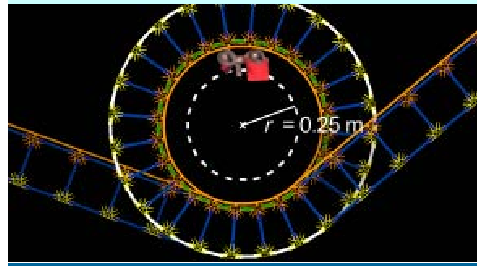
At top of loop:
Minimum speed at top of loop

$$mg = m\frac{v^2}{r}$$

$$\boxed{v = \sqrt{rg}}$$

m = mass
v = speed
r = radius
g = acceleration due to gravity

example 1



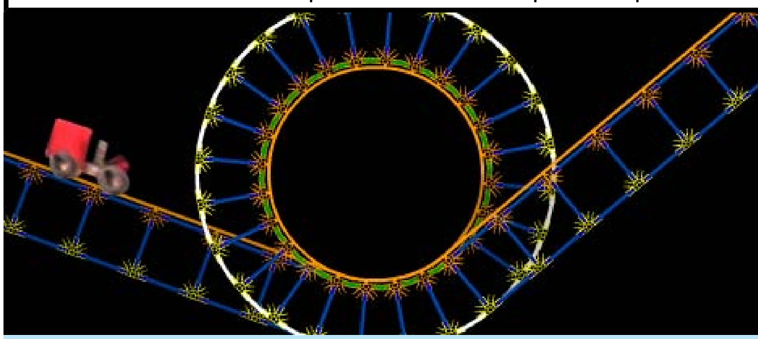
What is the minimum speed at the top for the toy to complete the loop?

$v = \sqrt{rg}$

$v = \sqrt{(0.25\text{ m})(9.80\text{ m/s}^2)}$

$v = 1.6$ m/s

9.12 - Interactive checkpoint: maximum loop-the-loop radius



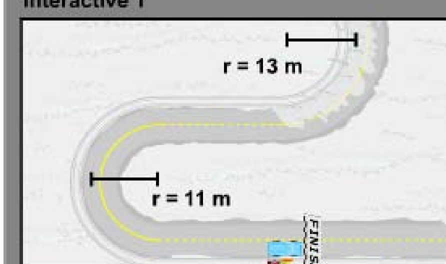
A roller coaster has a loop preceded by a ramp. A car is placed on the track so that its speed at the top of the loop will be 15.0 m/s. What is the maximum radius that the loop can have so that the car will not fall off the track?

Answer:

r = ______ m

9.13 - Interactive summary problem: race curves

interactive 1



Set speed of blue truck to win the race

In the simulation on the right, you are asked to race a truck on an S-shaped track against the computer. This time, the first curve is covered with snow and you are racing against a snowmobile. As you go around the track, the static friction between the tires of your truck and the snow or pavement provides the centripetal force. If you go too fast, you will exceed the maximum force of friction and your truck will leave the track. If you go as fast as you can without sliding, you will beat the snowmobile.

The snowmobile runs the entire race at its maximum speed. The blue truck negotiates each curve at a constant speed, but these speeds must be different for you to win the race. You set the speed of the blue truck on each curve. Straightaway sections are located at the start of the race and between the two curves. The simulation will automatically supply the acceleration you need on the straightaway sections.

The blue truck has a mass of 1,800 kg. The first curve is icy, and the coefficient of static friction of the truck on this curve is 0.51. (The snowmobile has a greater coefficient thanks to its snow-happy treads.) On the second curve, the coefficient of static friction is 0.84. The radius of the first curve is 13 m, and the second curve is 11 m. Set the speed of the blue truck on each curve as fast as it can go without sliding off the track, and you will win.

You set the speed in increments of 0.1 m/s in the simulation. If you need to round a value after your calculations, make sure you round down to the nearest 0.1 m/s. (If you round up, you will be exceeding the maximum safe speed.) Press GO to begin the race, and RESET if you need to try again.

If you have difficulty with this problem, you may want to review the section on static friction in a previous chapter and the section on centripetal acceleration in this chapter.

9.14 - Gotchas

A car is moving around a circular track at a constant speed of 20 km/h. This means its velocity is constant, as well. Wrong. The car's velocity changes because its direction changes as it moves.

Since an object moving in uniform circular motion is constantly changing direction, it is hard at any point in time to know the direction of its velocity and the direction of its acceleration. This is not true. The velocity vector is always tangent to the circle at the location of the object. Centripetal acceleration always points toward the center of the circle.

No force is required for an object to move in uniform circular motion. After all, its speed is constant. Yes, but its velocity is changing due to its change in direction, which means it is accelerating. By Newton's second law, this means there must be a net force causing this acceleration.

Centripetal force is another type of force. No, rather it is a way to describe what a force is "doing." The normal force, gravity, tension − each of these forces can be a centripetal force if it is causing an object to move in uniform circular motion.

9.15 - Summary

Uniform circular motion is movement in a circle at a constant speed. But while speed is constant in this type of motion, velocity is not. Since instantaneous velocity in uniform circular motion is always tangent to the circle, its direction changes as the object's position changes.

The period is the time it takes an object in uniform circular motion to complete one revolution of the circle.

Since the velocity of an object moving in uniform circular motion changes, it is accelerating. The acceleration due to its change in direction is called centripetal acceleration. For uniform circular motion, the acceleration vector has a constant magnitude and always points toward the center of the circle.

Newton's second law can be applied to an object in uniform circular motion. The net force causing centripetal acceleration is called a centripetal force. Like centripetal acceleration, it is directed toward the center of the circle.

Equations

$$T = \frac{2\pi r}{v}$$

$$a_c = \frac{v^2}{r}$$

$$F = m\frac{v^2}{r}$$

A centripetal force is not a new type of force; rather, it describes a role that is played by one or more forces in the situation, since there must be *some* force that is changing the velocity of the object. For example, the force of gravity keeps the Moon in a roughly circular orbit around the Earth, while the normal force of the road and the force of friction combine to keep a car in circular motion around a banked curve.

Conceptual Problems

C.1 A string's tension force supplies the centripetal force needed to keep a yo-yo whirling in a circle. (a) What force supplies the centripetal force keeping a satellite in uniform circular motion around the Earth? (b) What kinds of forces keep a roller coaster held to a looping track?

C.2 Health professionals use a device called a centrifuge to separate the different components of blood. If you allow a sample to sit long enough, Earth's gravity will cause it to separate on its own. This happens because the liquids and solids in blood have different densities. The denser solids sink to the bottom of a test tube, while less dense liquids rise to the top. To speed up the process of separation, a centrifuge spins blood sample tubes at high speeds in uniform circular motion. How does the spinning of the centrifuge speed up the separation process?

C.3 You move a roller coaster with a loop-the-loop from the Earth to your new amusement park on the Moon. (a) How does the minumum speed to complete the loop-the-loop compare between the Earth and the Moon? (b) A roller coaster car starts from rest on a hill that preceeds the loop. On the Earth, if the car is released on the hill from a height *h* above the bottom of the loop, it will have the minimum speed required to get around the loop. Is the release height required to just get around the same loop on the Moon greater than, less than, or equal to *h*?

(a) i. It is greater on the Moon
ii. It is less on the Moon
iii. There is no difference

(b) i. Greater
ii. Less
iii. Equal

C.4 Does an object moving in uniform circular motion have constant centripetal acceleration?

○ Yes ○ No

C.5 If a satellite in a circular orbit is accelerating toward the Earth, then why doesn't the satellite hit the Earth?

C.6 A father holds his three-year-old daughter's hands and swings her around in a circle, lifting her off her feet. Why is it harder for him to hold on the the faster he turns?

C.7 The hammer throw is a track-and-field event, popular in Scotland, in which a ball on a rope (the "hammer") is whirled around the thrower in a circle before being released. The goal is the send the ball as a projectile as far down the field as possible. At a track meet, the circle in which a ball is whirled by a hammer thrower is at a 45° angle to the ground. To achieve the longest distance, at what point in its tilted circular orbit should the thrower release the ball?

i. At its lowest point
ii. At its highest point
iii. Halfway as it moves from the lowest to the highest point
iv. Halfway as it moves from the highest to the lowest point

C.8 Two beads are tied to a string at different positions, and you swing the string around your head at a constant rate so that the beads move in uniform circular motion. Bead A is closer to your hand than bead B. Compare (a) the periods of A and B; (b) the speeds of A and B; (c) the centripetal accelerations of A and B.

(a) i. A and B have the same period
ii. A has a longer period
iii. B has a longer period

(b) i. A and B have the same speed
ii. A has a greater speed
iii. B has a greater speed

(c) i. A and B have same the acceleration
ii. A has a greater acceleration
iii. B has a greater acceleration

C.9 An object is moving at a constant speed around a circle. (a) In which of these cases does the magnitude of the centripetal acceleration of the object increase? (Assume all other factors are kept the same.) (b) In which case does the centripetal acceleration increase the most?

(a) ☐ The object's speed doubles
☐ The object's speed is halved
☐ The radius of the circle doubles
☐ The radius of the circle is halved

(b)
i. The object's speed doubles
ii. The object's speed is halved
iii. The radius of the circle doubles
iv. The radius of the circle is halved

C.10 Pretend that the Earth is rotating so fast that if you were standing at a fixed point on the equator, your weight would equal the centripetal force required to keep you in uniform circular motion around the Earth's center. If you stood on a scale at the equator, what would it read?

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Is the centripetal acceleration of the car higher when it is moving faster? (b) If the speed of the car remains constant, do the *x* and *y* components of the car's velocity change as the car goes around the track?

(a) ○ Yes ○ No
(b) ○ Yes ○ No

Section 2 - Period

2.1 Jupiter's distance from the Sun is 7.78×10^{11} meters and it takes 3.74×10^{8} seconds to complete one revolution of the Sun in its roughly circular orbit. What is Jupiter's speed?

__________ m/s

2.2 Saturn travels at an average speed of 9.66×10^{3} m/s around the Sun in a roughly circular orbit. Its distance from the Sun is 1.43×10^{12} m. How long (in seconds) is a "year" on Saturn?

__________ s

2.3 Mars travels at an average speed of 2.41×10^{4} m/s around the Sun, and takes 5.94×10^{7} s to complete one revolution. How far is Mars from the Sun?

__________ m

2.4 Long-playing vinyl records, still used by club DJs, are 12 inches in diameter and are played at 33 1/3 revolutions per minute. What is the speed (in m/s) of a point on the edge on such a record?

__________ m/s

Section 4 - Centripetal acceleration

4.1 A runner rounds a circular curve of radius 24.0 m at a constant speed of 5.25 m/s. What is the magnitude of the runner's centripetal acceleration?

__________ m/s^2

4.2 In a carnival ride, passengers are rotated at a constant speed in a seat at the end of a long horizontal arm. The arm is 8.30 m long, and the period of rotation is 4.00 s. (a) What is the magnitude of the centripetal acceleration experienced by a rider? (b) State the acceleration in "gee's," that is, as a multiple of the gravitational acceleration constant *g*.

(a) __________ m/s^2
(b) __________ g

4.3 Consider the radius of the Earth to be 6.38×10^{6} m. What is the magnitude of the centripetal acceleration experienced by a person (a) at the equator and (b) at the North Pole due to the Earth's rotation?

(a) __________ m/s^2
(b) __________ m/s^2

4.4 When tires are installed or reinstalled on a car, they are usually first balanced on a device that spins them to see if they wobble. A tire with a radius of 0.380 m is rotated on a tire balancing device at exactly 460 revolutions per minute. A small stone is embedded in the tread of the tire. What is the magnitude of the centripetal acceleration experienced by the stone?

__________ m/s^2

4.5 A toy airplane connected by a guideline to the top of a flagpole flies in a circle at a constant speed. If the plane takes 4.5 s to complete one loop, and the radius of the circular path is 11 m, what is the magnitude of the plane's centripetal acceleration?

__________ m/s^2

4.6 You tie a string to a rock and twirl it at a constant speed in a horizontal circle with a radius of 1.30 m, 2.10 m above the ground. The rock comes loose and travels as a projectile a horizontal distance of 9.40 m before striking the ground. (a) What was the magnitude of the centripetal acceleration of the rock when it was on the string? (b) What was its speed when it was on the string?

(a) __________ m/s^2
(b) __________ m/s

Section 5 - Interactive problems: racing in circles

5.1 Use the information given in the first interactive problem in this section to calculate the initial speed that keeps the blue car on the track and wins the race. For safety, round your answer **down** to the nearest 0.1 m/s. Test your answer using the simulation.

__________ m/s

5.2 Use the simulation in the second interactive problem in this section to calculate the initial speed that keeps the blue car on the track and wins the race. For safety, round your answer **down** to the nearest 0.1 m/s.

__________ m/s

Section 6 - Newton's second law and centripetal forces

6.1 An astronaut in training rides in a seat that is moved in uniform circular motion by a radial arm 5.10 meters long. If her speed is 15.0 m/s, what is the centripetal force on her in "G's," where one G equals her weight on the Earth?

__________ "G's"

6.2 A ball with mass 0.48 kg moves at a constant speed. A centripetal force of 23 N acts on the ball, causing it to move in a circle with radius 1.7 m. What is the speed of the ball?

__________ m/s

6.3 A bee loaded with pollen flies in a circular path at a constant speed of 3.20 m/s. If the mass of the bee is 133 mg and the radius of its path is 11.0 m, what is the magnitude of the centripetal force?

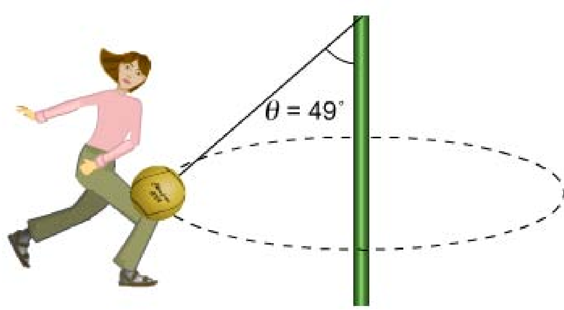
__________ N

6.4 Fifteen clowns are late to a party. They jump into their sporty coupe and start driving. Eventually they come to a level curve, with a radius of 27.5 meters. What is the top speed at which they can drive successfully around the curve? The coefficient of static friction between the car's tires and the road is 0.800.

__________ m/s

6.5 You are playing tetherball with a friend and hit the ball so that it begins to travel in a circular horizontal path. If the ball is 1.2 meters from the pole, has a speed of 3.7 m/s, a mass of 0.42 kilograms, and its (weightless) rope makes a 49° angle with the pole, find the tension force that the rope exerts on the ball just after you hit it.

__________ N



6.6 A car with mass 1600 kg drives around a flat circular track of radius 28.0 m. The coefficient of friction between the car tires and the track is 0.830. How fast can the car go around the track without losing traction?

__________ m/s

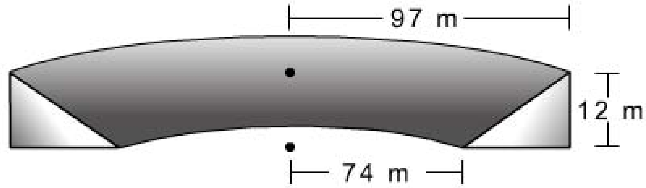
Section 7 - Sample problem: banked curves

7.1 A car with mass 2200 kg goes around a banked circular track in a path with radius 23 m. The track makes an angle of 15° with the horizontal. How fast should the car go around the track so that there will be no "sideways" friction force acting on the tires?

________ m/s

7.2 A car goes around the circular banked curve shown in the illustration at a speed of 21 m/s. At what radius does it travel if there is no frictional force between the car tires and the track?

________ m



7.3 What would the bank angle be for a circular racetrack with radius 120 m so that a car can go around the curve safely at a maximum of 21 m/s, without the help of frictional force to keep it on the road?

________ °

7.4 A jet of mass 9.7×10^4 kg makes a horizontal circular turn and "banks" or tilts its wings as it does so. The turn is 1200 m in radius and the jet has a constant speed of 160 m/s. The lift created by the plane's wings is perpendicular to the wing surface. (a) What is the magnitude of the lift force on the wings? (b) What is the angle of the lift force with respect to the horizontal?

(a) ________ N
(b) ________ °

Section 8 - Sample problem: centripetal force on a pendulum

8.1 You are at the carnival and decide to go on the swing ride. It is a high rotating platform from which swing seats hang like pendulums. As the platform begins to turn, your swing's chain does not stay perpendicular to the ground but angles out from the vertical. If your distance from the center of rotation is 8.50 meters and you go around once every 12.0 seconds, what angle does your swing make with the vertical?

________ °

8.2 You have lovingly restored a red 1964 Ford Falcon two-door sedan, complete with a pair of red fuzzy dice hanging from the rear-view mirror. When you go down a straight road at 26.0 m/s, the dice hang straight down, but when you enter a flat circular turn with radius 97.0 m at the same speed, they hang at an angle θ from the vertical. What is θ?

________ °

8.3 A tetherball is suspended on a 3.80 m rope from a tall pole. The ball is hit so that it travels in a horizontal circle around the pole with a constant speed of 5.60 m/s. What angle does the rope make with the pole?

________ °

Section 9 - Accelerating reference frames and fictitious forces

9.1 You hang a ball on a string from the ceiling inside a parked minivan. Then your friend starts up the van and drives forward at a constant acceleration while you measure the angle the string makes from its original position. If the angle is 9.00°, what is the magnitude of the van's acceleration?

________ m/s^2

Section 10 - Artificial gravity

10.1 A rotating space station has radius 1.31e+3 m, measured from the center of rotation to the outer deck where the crew lives. What should the period of rotation be if the crew is to feel that they weigh one-half their Earth weight?

________ s

Section 11 - Loop-the-loop

11.1 You are designing a roller coaster for a new amusement park on the Moon. You envision a loop with a radius of 35 meters in your track. (a) What minimum speed at the top will a car have to possess in order to successfully negotiate the loop? The gravitational acceleration on the Moon is 1.6 m/s^2. (b) What speed at the top of the loop would a car need to have if the same roller coaster was built on Earth?

(a) ________ m/s
(b) ________ m/s

11.2 A car drives over a hill that is shaped as a circular arc with radius 65.0 m. The car has a constant speed of 14.0 m/s and a mass of 423 kg. What is the magnitude of (a) the centripetal force on the car at the top of the hill and (b) the normal force exerted on the car by the road at this point?

(a) ________ N
(b) ________ N

11.3 A road has a hill with a top in the shape of a circular arc of radius 32.0 m. How fast can a car go over the top of the hill without losing contact with the ground?

________ m/s

11.4 A roller coaster consists of a hill followed by a loop. A car is released from rest on the hill from a height *h* above the bottom of the loop. If the car has the minimum speed at the top of the loop to avoid falling off, what is the ratio of *h* to *r*, the radius of the loop?

________ to 1

11.5 A roller coaster car does a vertical loop-the-loop with a radius of 23.0 meters. The centripetal force on a person in the car is 2.50 G's, which is 2.50 times her weight on Earth. How many times faster than the minimum velocity required to get around the loop is she going?

________ times the minimum velocity

Section 13 - Interactive summary problem: race curves

13.1 Use the simulation in the interactive problem in this section to calculate the speed for (a) the first curve and (b) the second curve that keeps the blue truck on the track and wins the race. For safety, round your answer **down** to the nearest 0.1 m/s.

(a) ________ m/s
(b) ________ m/s

Additional Problems

A.1 Rebecca goes on a popular ride at an amusement park that involves a cylinder rotating on a vertical axis, with a radius of 6.20 m. Rebecca stands inside, with her back against the wall of the cylinder. The ride rotates and presses Rebecca against the wall of the cylinder, and when she reaches a speed *v*, the floor of the ride drops so that she can no longer stand on it. The coefficient of static friction between the wall and Rebecca's clothing is 0.820. What should Rebecca's minimum speed *v* be when the floor is lowered, so that she does not fall?

________ m/s

A.2 The solar system moves around the galactic center in a roughly circular path at a radius of about 2.6×10^{20} m. The system's orbital speed around the center is 2.2×10^{5} m/s. (a) What is the period of this circular motion in years? (b) What is the sun's centripetal acceleration? (c) What is the magnitude of the centripetal force keeping the sun in this path given that the sun's mass is 1.99×10^{30} kg?

(a) ________ years
(b) ________ m/s^2
(c) ________ N

A.3 A soapbox racer rounds a curve banked at 28.0° with a radius of 21.0 meters. The tread on the racer's old tires is nearly gone so it can't provide much friction to keep the racer on track. How fast can the racer go around the curve without requiring friction to stay on track?

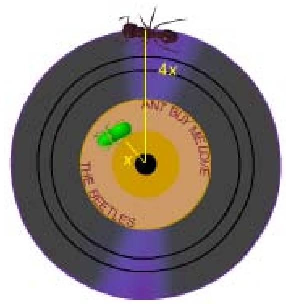
________ m/s

A.4 What period (in hours) would the Earth have to have for a person standing at the equator to experience a centripetal force equal to the gravitational force exerted on him by the Earth? When moving about, this person would be literally walking on air. The radius of the Earth is 6.38×10^{6} meters.

________ hours

A.5 An ant and a beetle are riding a spinning record on a record player. The ant stands on the outside edge of the record while the beetle stands on the record's label. The ant is four times as far from the center of rotation as the beetle. What is the ratio of the ant's centripetal acceleration to the beetle's?

__________ to 1



A.6 A dirty teddy bear is thrown into a top-loading washing machine. The cylinder in the machine rotates about a vertical axis. At what minimum speed does the wall of the cylinder have to move to stick the teddy to it so that he doesn't slide down? At this point, the frictional force acting upwards on the bear will balance his weight (this is the spin cycle, so the water has been drained and there is no buoyant force). The coefficient of static friction between a wet teddy bear and a washing machine wall is 0.500 and the washing machine cylinder has a radius of 0.310 meters.

__________ m/s

A.7 A car is traveling down a straight road at a constant 17 m/s. A tire on the car has a radius of 0.35 m. Imagine a dot on the edge of the tire. It will rotate in uniform circular motion with respect to the center of the tire, but the whole tire along with the rest of the car is moving linearly forward at the same time. What is the instantaneous velocity of the dot when it is (a) at the lowest point of its motion (on the road) and (b) at the highest point of its motion? (c) What is the average horizontal speed of the dot? (You do not need to do any calcuations to answer this last question.)

(a) __________ m/s
(b) __________ m/s
(c) __________ m/s

A.8 A car drives around two circular curves on two different roads. The two curves have the same radius of curvature. Coincidentally, the maximum speed that the car can drive through either of the turns is the same for both roads. The car drives through both turns at this speed. The first road is frictionless, but it is banked at 14 degrees off the horizontal. The other turn is flat. What is the coefficient of friction between the car tires and the road on the unbanked turn?

____

A.9 A relay satellite for satellite TV orbits the Earth above the equator at an altitude of 3.579×10^4 kilometers. This is the altitude required for a geosynchronous orbit − that is, for the satellite to remain above the same point on the Earth's equator as the Earth rotates. What is the speed of the satellite? The Earth has a radius of 6380 kilometers. (Note that a geosynchronous satellite's altitude is several times Earth's radius!)

__________ m/s

A.10 While on vacation you rent a moped and go cruising around a tropical island. If you take a 10.0° banked curve with a radius of 15.0 meters at 12.0 m/s, what is the minimum coefficient of static friction between your tires and the road to keep the tires from slipping as you negotiate the curve?

____

A.11 You are a traffic safety engineer in charge of determining safe speeds for roads. A particular banked curve has a radius of 11.0 meters and is banked at an angle of 8.00°. The coefficient of static friction between common tires and this road is 0.870. What is the maximum speed that a car can drive this curve? Use both the bank of the curve and the friction on the tires in determining your answer.

__________ m/s

10.0 - Introduction

If you feel as though you spend your life spinning around in circles, you may be pleased to know that an entire branch of physics is dedicated to studying that kind of motion. This chapter is for you! More seriously, this chapter discusses motion that consists of rotation about a fixed axis. This is called *pure rotational motion*. There are many examples of pure rotational motion: a spinning Ferris wheel, a roulette wheel, or a music CD are three instances of this type of motion.

In this chapter, you will learn about rotational displacement, rotational velocity, and rotational acceleration: the fundamental elements of what is called *rotational kinematics*.

You will also learn how to relate these quantities using equations quite similar to those used in the study of linear motion.

The simulation on the right features the "Angular Surge," an amusement park ride you will be asked to operate in order to gain insight into rotational kinematics. The ride has a rotating arm with a "rocket" where passengers sit. You can move the rocket closer to or farther from the center by setting the distance in the simulation. You can also change the rocket's period, which is the amount of time it takes to complete one revolution.

interactive 1

Change the rocket's linear speed and angular velocity

By changing these parameters, you affect two values you see displayed in gauges: the rocket's angular velocity and its linear speed. The rocket's angular velocity is the change per second in the angle of the ride's arm, measured from its initial position. Its units are radians per second. For instance, if the rocket completes one revolution in one second, its angular velocity is 2π radians $(360°)$ per second.

This simulation has no specific goal for you to achieve, although you may notice that you can definitely have an impact on the passengers! What you should observe is this: How do changes in the period affect the angular velocity? The linear speed? And how does a change in the distance from the center (the radius of the rocket's motion) affect those values, if at all? Can you determine how to maximize the linear speed of the rocket?

To run the ride, you start the simulation, set the values mentioned above, and press GO. You can change the settings while the ride is in motion.

10.1 - Angular position

Angular position: The amount of rotation from a reference position, described with a positive or negative angle.

When an object such as a bicycle wheel rotates about its axis, it is useful to describe this motion using the concept of angular position. Instead of being specified with a linear coordinate such as x, as linear position is, angular position is stated as an angle.

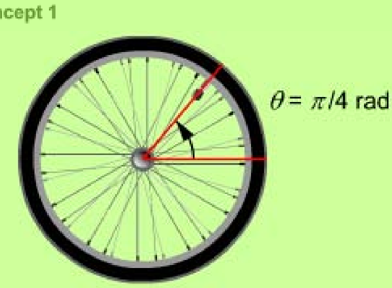
In Concept 1, we use the location of a bicycle wheel's valve to illustrate angular position. The valve starts at the 3 o'clock position (on the positive x axis), which is zero radians by convention. As the illustration shows, the wheel has rotated one-eighth of a turn, or $\pi/4$ radians $(45°)$, in a counterclockwise direction away from the reference position. In other words, angular position is measured from the positive x axis.

Note that this description of the wheel's position used radians, not degrees; this is because radians are typically used to describe angular position. The two lines we use to measure the angle radiate from the point about which the wheel rotates.

The *axis of rotation* is a line also used to describe an object's rotation. It passes through the wheel's center, since the wheel rotates about that point, and it is perpendicular to the wheel. The axis is assumed to be stationary, and the wheel is assumed to be rigid and to maintain a constant shape. Analyzing an object that changes shape as it rotates, such as a piece of soft clay, is beyond the scope of this textbook. We are concerned with the wheel's rotational motion here: its motion around a fixed axis. Its linear motion when moving along the ground is another topic.

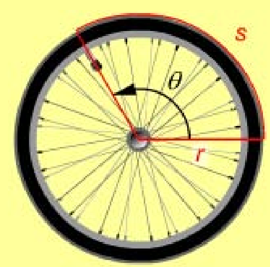
As mentioned, angular position is typically measured with *radians* (rad) instead of degrees. The formula that defines the radian measure of an angle is shown in Equation 1. The angle in radians equals the arc length s divided by the radius r. As you may recall, 2π radians equals one revolution around a circle, or $360°$. One radian equals about $57.3°$. To convert radians to degrees, multiply by the conversion factor $360°/2\pi$. To convert degrees to radians, multiply by the reciprocal: $2\pi/360°$. The Greek letter θ

concept 1



Angular position
Rotation from 3 o'clock position
· Counterclockwise rotation: positive
· Clockwise rotation: negative
Units are radians

equation 1



Radian measure

(theta) is used to represent angular position.

The angular position of zero radians is defined to be at 3 o'clock, which is to say along a horizontal line pointing to the right. Let's now consider what happens when the wheel rotates a quarter turn **counterclockwise**, moving the valve from the 3 o'clock position to 12 o'clock. A quarter turn is $\pi/2$ rad (or $90°$). The valve's angular position when it moves a quarter turn counterclockwise is $\pi/2$ rad. By convention, angular position **increases** with counterclockwise motion.

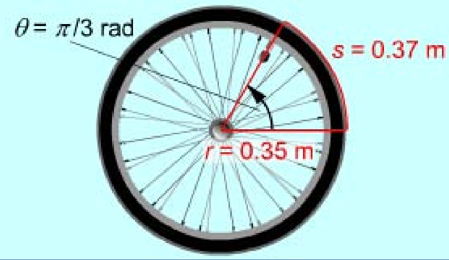
The valve can be placed in the same angular position, $\pi/2$ rad, by rotating the wheel in the other direction, by rotating it **clockwise** three quarters of a turn. By convention, angular position **decreases** with clockwise motion, so this rotation would be described as an angular position of $-3\pi/2$ rad.

An angular position can be greater than 2π rad. An angular position of 3π rad represents one and a half counterclockwise revolutions. The valve would be at 9 o'clock in that position.

$$\theta = \frac{s}{r}$$

θ = angle in radians
s = arc length
r = radius

example 1



What is the arc length?

$$\theta = \frac{s}{r}$$

$s = r\theta = (0.35\text{ m})(\pi/3\text{ rad})$
$s = 0.37\text{ m}$

10.2 - Angular displacement

Angular displacement: Change in angular position.

In Concept 1 you see a pizza topped with a single mushroom (we are not going back to that pizzeria!). We use a mushroom to make the rotational motion of the pizza easier to see. As the pizza rotates, its angular position changes. This change in angular position is called angular displacement.

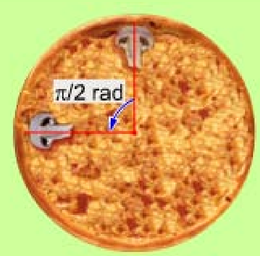
To calculate angular displacement, you subtract the initial angular position from the final position. For instance, the mushroom in the Equation illustration moves from $\pi/2$ rad to π rad, a displacement of $\pi/2$ rad. As you can see in this example, angular displacement in the counterclockwise direction is positive.

Revolution is a common term in the study of rotational motion. It means one complete rotational cycle, with the object starting and returning to the same position. One counterclockwise revolution equals 2π radians of angular displacement.

The angular displacement is the **total** angle "swept out" during rotational motion from an initial to a final position. If the pizza turns counterclockwise three complete revolutions, its angular displacement is 6π radians.

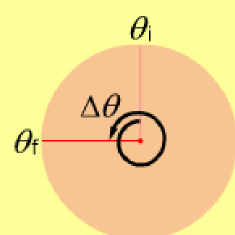
The definition of angular displacement resembles that of linear displacement. However, the discussion above points out a difference. A mushroom that makes a complete revolution has an angular displacement of 2π rad. On the other hand, its linear displacement equals zero, since it starts and stops at the same point.

concept 1



Angular displacement
Change in angular position
Counterclockwise rotation is positive

equation 1



Angular displacement

$$\Delta\theta = \theta_f - \theta_i$$

$\Delta\theta$ = angular displacement
θ_f = final angular position
θ_i = initial angular position

Units: radians (rad)

example 1

At 12:10, the initial angular position of the minute hand is $\pi/6$. After 15 minutes have passed, what is the minute hand's angular displacement?

$$\Delta\theta = \theta_f - \theta_i$$

$$\Delta\theta = \left(-\frac{\pi}{3}\text{ rad}\right) - \left(\frac{\pi}{6}\text{ rad}\right)$$

$$\Delta\theta = -\frac{\pi}{2}\text{ rad}$$

10.3 - Angular velocity

Angular velocity: Angular displacement per unit time.

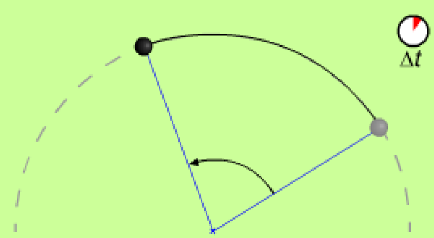
In Concept 1, a ball attached to a string is shown moving counterclockwise around a circle. Every four seconds, it completes one revolution of the circle. Its angular velocity is the angular displacement 2π radians (one revolution) divided by four seconds, or $\pi/2$ rad/s. The Greek letter ω (omega) represents angular velocity.

As is the case with linear velocity, angular velocity can be discussed in terms of average and instantaneous velocity. *Average angular velocity* equals the total angular displacement divided by the elapsed time. This is shown in the first equation in Equation 1.

Instantaneous angular velocity refers to the angular velocity at a precise moment in time. It equals the limit of the average velocity as the increment of time approaches zero. This is shown in the second equation in Equation 1.

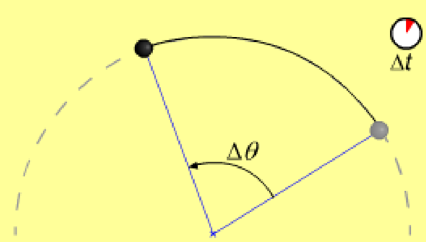
The sign of angular velocity follows that of angular displacement: positive for counterclockwise rotation and negative for clockwise rotation. The magnitude (absolute value) of angular velocity is *angular speed*.

concept 1



Angular velocity
Angular displacement per unit time

equation 1



Angular velocity

$$\bar{\omega} = \frac{\Delta\theta}{\Delta t}$$

$\bar{\omega}$ = average angular velocity

$$\omega = \lim_{\Delta t \to 0} \frac{\Delta\theta}{\Delta t}$$

ω = instantaneous angular velocity
$\Delta\theta$ = angular displacement
Δt = elapsed time
Units: rad/s

example 1

The motorcycle rider takes two seconds for a counterclockwise lap around the track. What is her average angular velocity?

$$\bar{\omega} = \frac{\Delta\theta}{\Delta t}$$

$$\bar{\omega} = \frac{2\pi \text{ rad}}{2 \text{ s}}$$

$$\bar{\omega} = \pi \text{ rad/s}$$

10.4 - Angular acceleration

Angular acceleration: The change in angular velocity per unit time.

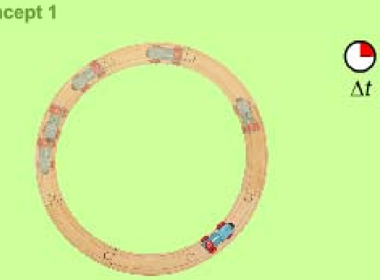
By now, you might be experiencing a little *déjà vu* in this realm of angular motion. Angular velocity equals angular displacement per unit time, but if you drop the word "angular" you are stating that velocity equals displacement per unit time, an equation that should be familiar to you from your study of linear motion.

So it is with angular acceleration. Angular acceleration equals the change in angular velocity divided by the elapsed time. The toy train shown in Concept 1 is experiencing angular acceleration. This is reflected in the increasing separation between the images you see. Its angular **velocity** is becoming increasingly negative since it is moving in the clockwise direction. It is moving faster and faster in the negative angular direction.

Average angular acceleration equals the change in angular velocity divided by the elapsed time. The *instantaneous angular acceleration* equals the limit of this ratio as the increment of time approaches zero. These two equations are shown in Equation 1 to the right. The Greek letter α (alpha) is used to represent angular acceleration.

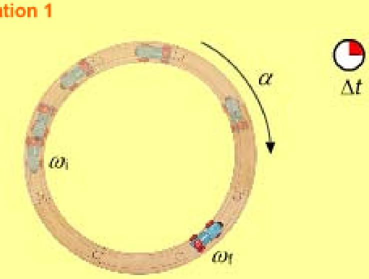
With rotational kinematics, we often pose problems in which the angular acceleration is constant; this helps to simplify the mathematics involved in solving problems. We made similar use of constant acceleration for the same reason in the linear motion chapter.

concept 1



Angular acceleration
Change in angular velocity per unit time

equation 1



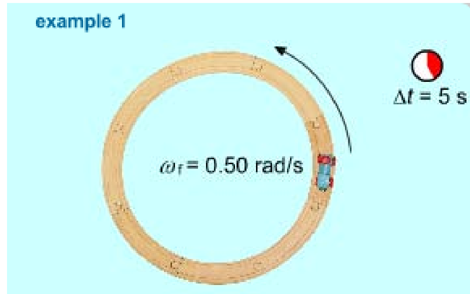
Angular acceleration

$$\bar{\alpha} = \frac{\Delta\omega}{\Delta t}$$

$\bar{\alpha}$ = average acceleration

$$\alpha = \lim_{\Delta t \to 0} \frac{\Delta \omega}{\Delta t}$$

α = instantaneous angular acceleration
ω = angular velocity
Δt = elapsed time
Units: rad/s^2



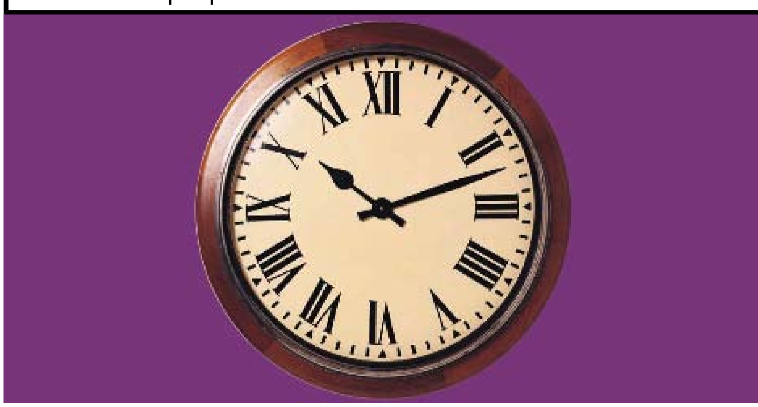
The toy train starts from rest and reaches the angular velocity shown in 5.0 seconds. What is its average angular acceleration?

$$\bar{\alpha} = \frac{\Delta \omega}{\Delta t}$$

$$\bar{\alpha} = \frac{0.50 \text{ rad/s} - 0.00 \text{ rad/s}}{5.0 \text{ s}}$$

$$\bar{\alpha} = 0.10 \text{ rad/s}^2$$

10.5 - Sample problem: a clock



Over the course of 1.00 hour, what is (a) the angular displacement, (b) the angular velocity and (c) the angular acceleration of the minute hand?

Think about the movement of the minute hand over the course of an hour. Be sure to consider the direction!

Variables

elapsed time	Δt = 1.00 h
angular displacement	$\Delta\theta$
angular velocity	ω
angular acceleration	α

What is the strategy?

1. Calculate the angular displacement.
2. Convert the elapsed time to seconds.
3. Use the angular displacement and time to determine the angular velocity and angular acceleration.

Physics principles and equations

Definition of angular velocity

$$\omega = \frac{\Delta\theta}{\Delta t}$$

Definition of angular acceleration

$$\alpha = \frac{\Delta\omega}{\Delta t}$$

Step-by-step solution

We start by calculating the angular displacement of the minute hand over 1.00 hour. We then calculate the angular velocity.

Step	Reason
1. $\Delta\theta = -2\pi$ rad	minute hand travels clockwise one revolution
2. $\Delta t = 1.00\text{ h}\left(\frac{3600\text{ s}}{1.00\text{ h}}\right) = 3600\text{ s}$	convert to seconds
3. $\omega = \frac{\Delta\theta}{\Delta t}$	definition of angular velocity
4. $\omega = \frac{-2\pi\text{ rad}}{3600\text{ s}}$	substitute
5. $\omega = -1.75\times10^{-3}$ rad/s	evaluate

The angular displacement is calculated in step 1, and the angular velocity in step 5. Since the angular velocity is constant, the angular acceleration is zero.

10.6 - Interactive checkpoint: a potter's wheel

At a particular instant, a potter's wheel rotates clockwise at 12.0 rad/s; 2.50 seconds later, it rotates at 8.50 rad/s clockwise. Find its average angular acceleration during the elapsed time.

Answer:

$\bar{\alpha} =$ ______ rad/s^2

10.7 - Equations for rotational motion with constant acceleration

$$\omega_f = \omega_i + \alpha t$$

$$\Delta\theta = \omega_i t + \frac{1}{2}\alpha t^2$$

$$\omega_f^2 = \omega_i^2 + 2\alpha\Delta\theta$$

$$\Delta\theta = \frac{1}{2}(\omega_i + \omega_f)t$$

Δθ = angular displacement, ω = angular velocity, α = angular acceleration, t = elapsed time

equation 1

Rotational motion equations

To solve problems

· Identify known and unknown values

· Choose equation that includes those values

Above, you see a table of equations for rotational motion. These equations resemble those for linear motion, except that they now are defined for angular displacement, angular velocity and angular acceleration instead of linear displacement, velocity and acceleration. As with the linear motion equations, these equations hold true when there is constant acceleration. We also show these equations below along with their linear counterparts.

To apply the equations in physics problems, the first step is to identify the known values and which values are being asked for. Sketching a diagram of the situation may help you with this.

The next step is to find an equation that includes both the known and the unknown (asked-for) values. Your goal is to find an equation, if possible, that has only one unknown value: the one you want to find.

When applying the rotational equations, remember that positive displacement and velocity represent counterclockwise motion, and negative displacement and velocity indicate clockwise motion.

Let's now work an example problem. Imagine you have just turned on the blender shown on the right. You let it run for 5.0 seconds. During this time period its blade has a constant angular acceleration of 44 radians per second squared. What is the angular displacement of the blade during this time?

This problem implicitly tells you that the initial angular velocity is zero, since the blender has just been turned on. The second equation above includes time, initial angular velocity and acceleration. It also contains the value you seek to calculate: the angular displacement. This makes it the right equation to use. It does not include the value for final angular velocity, which is fine because you are not told that value, nor are you asked to calculate it.

The details of the calculation appear on the right. The angular displacement is 550 radians. Because the value is positive, the motion is counterclockwise.

example 1

The blender is turned on and runs for 5.0 seconds with a constant angular acceleration of 44 rad/s². What is the angular displacement of a blade?

$\Delta\theta = \omega_i t + \frac{1}{2}\alpha t^2$

$\omega_i = 0$ rad/s

$\Delta\theta = 0$ rad $+ \frac{1}{2}(44 \text{ rad/s}^2)(5.0 \text{ s})^2$

$\Delta\theta = 0 + (22)(25)$ rad

$\Delta\theta = 550$ rad

Here is a table of the rotational motion variables and the equations that relate them, along with their linear counterparts.

	linear	rotational
position	x	θ
displacement	Δx	$\Delta\theta$
velocity	$v = \Delta x/\Delta t$	$\omega = \Delta\theta/\Delta t$
acceleration	$a = \Delta v/\Delta t$	$\alpha = \Delta\omega/\Delta t$
	$v_f = v_i + at$	$\omega_f = \omega_i + \alpha t$
	$\Delta x = v_i t + \frac{1}{2}at^2$	$\Delta\theta = \omega_i t + \frac{1}{2}\alpha t^2$
	$v_f^2 = v_i^2 + 2a\Delta x$	$\omega_f^2 = \omega_i^2 + 2\alpha\Delta\theta$
	$\Delta x = \frac{1}{2}(v_i + v_f)t$	$\Delta\theta = \frac{1}{2}(\omega_i + \omega_f)t$

10.8 - Sample problem: a carousel

The carousel accelerates from rest for two revolutions at a constant angular acceleration of 0.11 rad/s². What is its final angular velocity?

Variables

Since the carousel is starting up, the initial angular velocity must be zero. The angular displacement is given in revolutions, which must be converted to radians.

initial angular velocity	$\omega_i = 0$ rad/s
final angular velocity	ω_f
angular acceleration	$\alpha = 0.11$ rad/s²
angular displacement	$\Delta\theta = (2.0)(2\pi$ rad$)$

What is the strategy?

1. Identify the known and unknown values and choose an equation.
2. Substitute the known values and solve the equation for the final angular velocity.

Physics principles and equations

For the known and unknown values in this problem, the appropriate equation is

$$\omega_f^2 = \omega_i^2 + 2\alpha\Delta\theta$$

This includes the values stated above and the value we are asked to determine.

Step-by-step solution

Step	Reason
1. $\omega_f^2 = \omega_i^2 + 2\alpha\Delta\theta$	rotational motion equation
2. $\omega_f^2 = (0.0 \text{ rad/s})^2 + 2(0.11 \text{ rad/s}^2)(2.0)(2\pi \text{ rad})$	enter values
3. $\omega_f^2 = 2.76 \text{ rad}^2/\text{s}^2$	evaluate
4. $\omega_f = 1.7 \text{ rad/s}$	square root

10.9 - Interactive checkpoint: roulette

A roulette wheel is spun counter-clockwise with an initial angular velocity of 15.0 rad/s. If it slows down with a constant angular acceleration of −0.200 rad/s^2 until it stops, what is the total angular displacement of the wheel?

Answer:

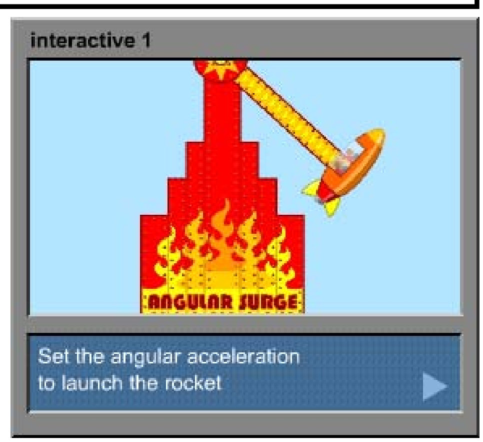
$\Delta\theta =$ ______ rad

10.10 - Interactive problem: launch the rocket

Lucky you! You just landed a summer job operating the Angular Surge ride at a local amusement park. As shown on the right, the ride has a rotating arm with a rocket for transporting riders. The rocket is fixed to the end of the arm, and you control the angular acceleration of the ride for the first revolution. Your goal is to set this constant acceleration so that, after exactly one revolution, the ride has an angular velocity of 1.64 rad/s. If you set this value correctly, the rocket will take off. It must be moving at precisely this angular velocity to have the correct amount of energy to safely launch.

The angular acceleration you specify will be applied for exactly one revolution and then the rocket arm will maintain its velocity. Calculate the angular acceleration, enter the value and press GO to see the results. If you do not succeed in launching the rocket, press RESET to try again.

If you have difficulty with this problem, review the rotational motion equations and make sure you have chosen the appropriate one to solve the problem.



Tangential velocity: The instantaneous linear velocity of a point on a rotating object.

Concepts such as angular displacement and angular velocity are useful tools for analyzing rotational motion. However, they do not provide the complete picture. Consider the salt and pepper shakers rotating on the lazy Susan shown to the right. The containers have the same angular velocity because they are on the same rotating surface and complete a revolution in the same amount of time.

However, at any instant, they have different **linear** speeds and velocities. Why? They are located at different distances from the axis of rotation (the center of the lazy Susan), which means they move along circular paths with different radii. The circular path of the outer shaker is longer, so it moves farther than the inner one in the same amount of time. At any instant, its linear speed is greater. Because the direction of motion of an object moving in a circle is always tangent to the circle, the object's linear velocity is called its tangential velocity.

To reinforce the distinction between linear and angular velocity, consider what happens if you decide to run around a track. Let's say you are asked to run one lap around a circular track in one minute flat. Your angular velocity is 2π radians per minute.

Could you do this if the track had a radius of 10 meters? The answer is yes. The circumference of that track is $2\pi r$, which equals approximately 63 meters. Your pace would be that distance divided by 60 seconds, which works out to an easy stroll of about 1.05 m/s (3.78 km/h).

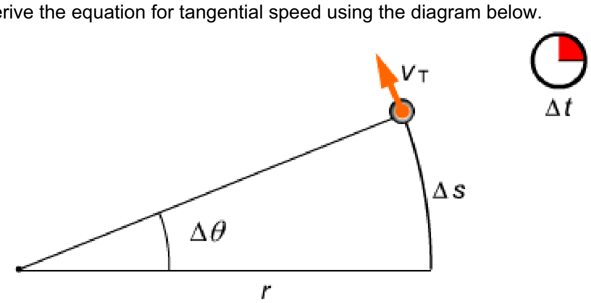
What if the track had a radius of 100 meters? In this case, the one-minute accomplishment would require the speed of a world-class sprinter capable of averaging more than 10 m/s. (If the math ran right past you, note that we are again multiplying the radius by 2π to calculate the circumference and dividing by 60 seconds to calculate the tangential velocity.) Even though the angular velocity is the same in both cases, 2π radians per minute, the tangential speed changes with the radius.

As you see in Equation 1, tangential speed equals the product of the distance to the axis of rotation, r, and the angular velocity, ω. The units for tangential velocity are meters per second. The direction of the velocity is always tangent to the path of the object.

Confirming the direction of tangential velocity can be accomplished using an easy home experiment. Let's say you put a dish on a lazy Susan and then spin the lazy Susan faster and faster. Initially, the dish moves in a circle, constrained by static friction. At some point, though, it will fly off. The dish will always depart in a straight line, tangent to the circle at its point of departure.

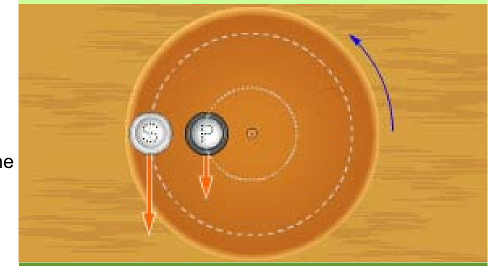
The tangential speed equation can also be used to restate the equation for centripetal acceleration in terms of angular velocity. Centripetal acceleration equals v^2/r. Since $v = r\omega$, centripetal acceleration also equals $\omega^2 r$.

We derive the equation for tangential speed using the diagram below.



To understand the derivation, you must recall that the arc length Δs (the distance along the circular path) equals the angular displacement $\Delta\theta$ in radians times the radius r. Also recall that the instantaneous speed v_T equals the displacement divided by the elapsed time for a very small increment of time.

concept 1



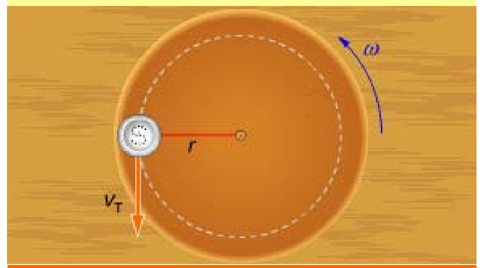
Tangential velocity

Linear velocity at an instant

· Magnitude: magnitude of linear velocity

· Direction: tangent to circle

equation 1



Tangential velocity

$$v_T = r\omega$$

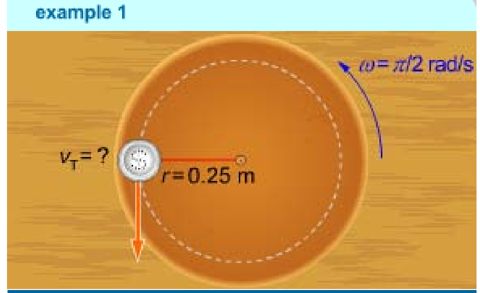
v_T = tangential speed

r = distance to axis

ω = angular velocity

Direction: tangent to circle

example 1



At the instant shown, what is the salt shaker's tangential velocity?

$v_T = r\omega$

$v_T = (0.25 \text{ m})(\pi/2 \text{ rad/s})$

$\mathbf{v}_T = 0.39$ m/s, pointing down

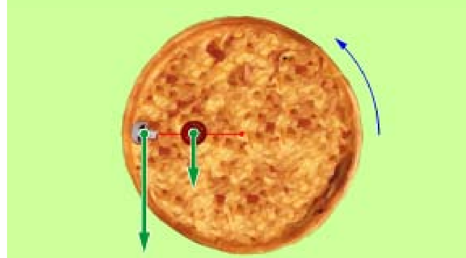
Combining these two facts, and the definition of angular velocity, yields the equation for tangential speed.

Step	Reason
1. $v_T = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}$	definition of instantaneous velocity
2. $\Delta s = r\Delta\theta$	definition of radian measure
3. $v_T = \lim_{\Delta t \to 0} \frac{r\Delta\theta}{\Delta t} = r\left(\lim_{\Delta t \to 0} \frac{\Delta\theta}{\Delta t}\right)$	substitute equation 2 into equation 1
4. $v_T = r\omega$	definition of angular velocity

10.12 - Tangential acceleration

Tangential acceleration: A vector tangent to the circular path whose magnitude is the rate of change of tangential speed.

concept 1



Tangential acceleration
Rate of change of tangential speed
Increases with distance from center
Direction of vector is tangent to circle

As discussed earlier, an object moving in a circle at a constant speed is accelerating because its direction is constantly changing. This is called centripetal acceleration.

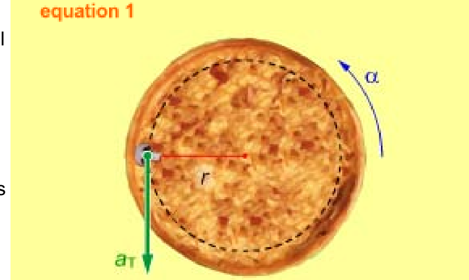
Now consider the mushroom on the pizza to the right. Let's say the pizza has a positive angular acceleration. Since it is rotating faster and faster, its angular velocity is increasing. Since tangential speed is the product of the radius and the angular velocity, the magnitude of its tangential velocity is also increasing.

The magnitude of the tangential acceleration vector equals the rate of change of tangential speed. The tangential acceleration vector is always parallel to the linear velocity vector. When the object is speeding up, it points in the same direction as the tangential velocity vector; when the object is slowing down, tangential acceleration points in the opposite direction.

Since the centripetal acceleration vector always points toward the center, the centripetal and tangential acceleration vectors are perpendicular to each other. An object's overall acceleration is the sum of the two vectors. To put it another way: The centripetal and tangential acceleration are perpendicular components of the object's overall acceleration.

Like tangential velocity, tangential acceleration increases with the distance from the axis of rotation. Consider again the pizza and its toppings in Concept 1. Imagine that the pizza started stationary and it now has positive angular acceleration. Since tangential velocity is proportional to radius, at any moment in time the mushroom near the outer edge of the pizza has greater tangential velocity than the piece of pepperoni closer to the center. Since the mushroom's change in tangential velocity is greater, it must have accelerated at a greater rate.

equation 1



Tangential acceleration

$$a_T = r\alpha$$

a_T = tangential acceleration
r = distance to axis
α = angular acceleration
Direction: tangent to circle

Tangential acceleration can be calculated as the product of the radius and the angular acceleration. This relationship is stated in Equation 1. The units for tangential acceleration are meters per second squared, the same as for linear acceleration. Note that it only makes sense to calculate the tangential acceleration for an object (or really a point) on the pizza. You cannot speak of the tangential acceleration of the entire pizza because it includes points that are at different distances from its center and have different rates of tangential acceleration.

Because it is easy to confuse angular and linear motion, we will now review a few fundamental relationships.

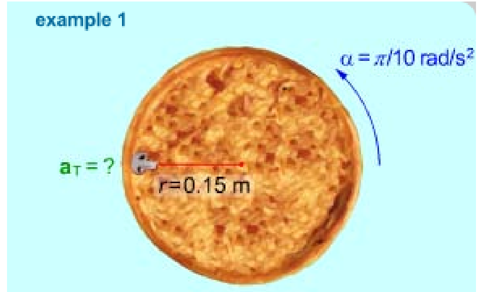
An object rotating at a constant angular velocity has zero angular acceleration and zero tangential acceleration. An example of this is a car driving around a circular track at a constant speed, perhaps at 100 km/hr. This means the car completes a lap at a constant rate, so its angular velocity is constant. A constant angular velocity means zero angular acceleration. Since the angular acceleration is zero, so is the tangential acceleration.

In contrast, the car's linear (or tangential) velocity is changing since it changes direction as it moves along the circular path. This accounts for the car's centripetal acceleration, which equals its speed squared divided by the radius of the track. The direction of centripetal acceleration is always toward the center of the circle.

Now imagine that the car speeds up as it circles the track. It now completes a lap more quickly, so its angular velocity is increasing, which means it has positive angular acceleration (when it is moving counterclockwise; it is negative in the other direction). The car now has tangential acceleration (its linear speed is changing), and this can be calculated by multiplying its angular acceleration by the track's radius.

The equation for tangential acceleration is derived below from the equations for tangential velocity and angular acceleration. We begin with the basic definition of linear acceleration and substitute the tangential velocity equation. The result is an expression which contains the definition of angular acceleration. We replace this expression with α, angular acceleration, which yields the equation we desire.

Step	Reason
1. $a_T = \lim_{\Delta t \to 0} \frac{\Delta v_T}{\Delta t}$	definition of linear acceleration
2. $\Delta v_T = r\Delta\omega$	tangential velocity equation
3. $a_T = \lim_{\Delta t \to 0} \frac{r\Delta\omega}{\Delta t} = r\left(\lim_{\Delta t \to 0} \frac{\Delta\omega}{\Delta t}\right)$	substitute equation 2 into equation 1
4. $a_T = r\alpha$	definition of angular acceleration



What is the tangential acceleration of the mushroom slice at this instant?

$a_T = r\alpha$

$a_T = (\pi/10 \text{ rad/s}^2)(0.15 \text{ m})$

$\mathbf{a}_T = 0.047 \text{ m/s}^2$, pointing down

10.13 - Interactive checkpoint: a marching band

The performers in a marching band move in straight rows, maintaining constant side-to-side spacing between them. Each row sweeps 90° through a circular arc when the band turns a corner. The radii of the paths followed by the marchers at the inner and outer ends of a row are 1.50 m and 7.50 m. If the innermost marcher in a row moves at 0.350 m/s, what is the speed of the outermost marcher?

Answer:

v_{out} = ______ m/s

10.14 - Tangential and centripetal acceleration

In Concept 1, a toy train is shown going around a circular track at steadily increasing speed. How can we calculate its overall acceleration at any moment?

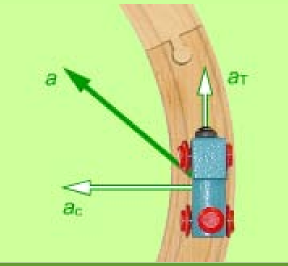
The train has both centripetal and tangential acceleration. The overall acceleration can be broken into these two components.

The acceleration perpendicular to the direction of motion, directed toward the center of the circle, is the centripetal acceleration. Its magnitude at any instant is calculated using the equation for centripetal acceleration from a previous chapter: speed squared divided by the radius.

The acceleration parallel to the velocity vector is the tangential acceleration, which is perpendicular to the centripetal acceleration. Since the train is increasing in speed, it has non-zero tangential acceleration. (This is not uniform circular motion.)

The overall acceleration equals the vector sum of the centripetal and tangential accelerations. The two vectors are perpendicular, so they form two legs of a right triangle. The Pythagorean theorem can be used to calculate the magnitude of the overall acceleration, as the first formula in Equation 1 shows. The direction of the overall acceleration, measured from the centripetal acceleration vector (or the radius line), can be calculated using trigonometry. You see that formula in Equation 1 as well.

concept 1



Combining centripetal and tangential acceleration

Centripetal acceleration

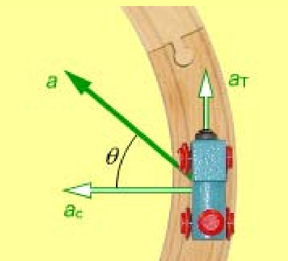
· Points toward center of circular path

Tangential acceleration

· Reflects speeding or slowing

Overall acceleration is the vector sum

equation 1



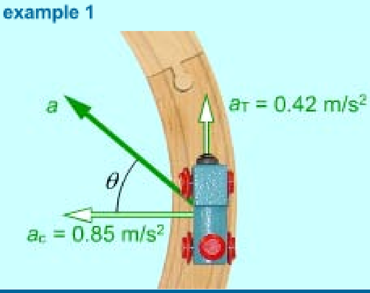
Combining centripetal and tangential acceleration

$$a = \sqrt{a_c^2 + a_T^2}$$

$$\theta = \arctan\,(a_T/a_c)$$

a = overall acceleration,
a_c = centripetal acceleration,
a_T = tangential acceleration,
θ = angle of a relative to a_c

example 1



What are the magnitude and direction of the overall acceleration?

$$a = \sqrt{a_c^2 + a_T^2}$$

$$a = \sqrt{(0.85\ \text{m/s}^2)^2 + (0.42\ \text{m/s}^2)^2}$$

$a = 0.95\ \text{m/s}^2$

$\theta = \arctan\,(a_T/a_c)$

$\theta = \arctan(0.42/0.85) = 26°$

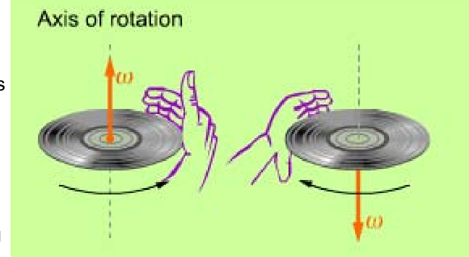
10.15 - Vectors and angular motion

Although we have not stressed this fact, angular velocity and angular acceleration are both vectors. In this section, we discuss the direction in which they point, using the *right-hand rule* to determine their direction. To apply this rule to angular velocity, curl your right hand around the axis of rotation, wrapping your fingers in the direction of the motion. This is illustrated to the right, where the hand wraps around the axis that passes through the center of the record. Your thumb then points in the direction of the angular velocity vector, which lies along the axis of rotation.

The direction of the angular acceleration vector depends on whether the object in question is speeding up or slowing down. When an object speeds up, the angular acceleration vector points in the same direction as the angular velocity vector, reflecting the change in the velocity vector. When an object slows down, the angular acceleration vector points in the direction **opposite** to the angular velocity vector, again reflecting the change in the angular velocity vector.

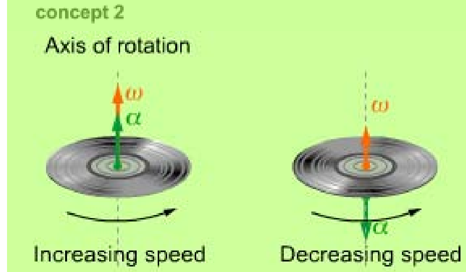
You may have noticed that we have not mentioned angular displacement. This is because it is not treated as a vector. Understanding why this is so requires a discussion of rotational motion outside the scope of this textbook.

concept 1



Angular velocity vector
Along axis of rotation
Magnitude proportional to angular speed
Direction determined by right-hand rule

concept 2



Angular acceleration vector
Speeding up: same direction as angular velocity vector
Slowing down: opposite direction

example 1

The motorcycle rider speeds up as she starts her ride. What is the direction of the angular velocity vector? The angular acceleration vector?

Angular velocity vector: up

Angular acceleration vector: up

10.16 - Interactive summary problem: 11.6 seconds to liftoff

You are again operating the Angular Surge ride at a local amusement park. The ride begins with the arm in the launch position for the rocket. The motor starts the ride by providing a constant positive angular acceleration for the first 11.6 seconds.

The ride has a rocket on a rotating arm, and you can control the arm's angular acceleration. You can also control the distance of the rocket from the axis of rotation.

Your goal is to set both these values so that 11.6 seconds after startup, the rocket has completed one or more complete revolutions **and** has a tangential velocity of 13.0 m/s. If you do this correctly, the rocket will blast off. You can position the rocket from four to 10 meters from the center, in increments of one-half meter.

The rocket can complete one revolution, or multiple complete revolutions, as long as it returns to its initial position in 11.6 seconds. Here is one way to solve the problem: Start by calculating the angular acceleration that would be needed to complete one revolution in 11.6 seconds, to the nearest hundredth rad/s^2, and enter it in the space provided. From this, you can calculate the final angular velocity of the rocket. Can you set the radius of the rocket so that its tangential velocity is 13.0 m/s? If not, try again, using two complete revolutions.

When you have determined the angular acceleration you want to use and the radius required, set these values, then press GO. If you are correct, at 11.6 seconds the rocket will take off.

If you have difficulty with this problem, review the sections on the rotational motion equations and tangential velocity, and remember that more than one revolution may be necessary.

10.17 - Gotchas

A potter's wheel rotates. A location farther from the axis will have a greater angular velocity than one closer to the axis. Wrong. They all have the same angular displacement over time, which means they have the same angular velocity, as well. In contrast, they do have different **linear** (tangential) velocities.

A point on a wheel rotates from 12 o'clock to 3 o'clock, so its angular displacement is 90 degrees, correct? No. This would be one definite error and one "units police" error. The displacement is negative because clockwise motion is negative. And, using radians is preferable and sometimes essential in the study of angular motion, so the angular displacement should be stated as $-\pi/2$ radians.

10.18 - Summary

Rotational kinematics applies many of the ideas of linear motion to rotational motion.

Angular position is described by an angle θ, measured from the positive *x* axis. Radians are the typical units.

Angular displacement is a change $\Delta\theta$ in angular position. By convention, the counterclockwise direction is positive.

Angular velocity is the angular displacement per unit time. It is represented by ω and has units of radians per second.

Angular acceleration is the change in angular velocity per unit time. It is represented by α and has units of radians per second squared.

As with linear motion, physicists define instantaneous and average angular velocity and angular acceleration. Instantaneous and average are defined in ways analogous to those used in the study of linear motion.

The equations for rotational motion are the same as those for linear motion except that displacement, velocity and constant acceleration are replaced by angular displacement, angular velocity and angular acceleration. For the equations to hold true, the angular acceleration must be constant.

The linear velocity of a point on a rotating object is called its tangential velocity, because it is always directed tangent to its circular path. Any two points on a rigid rotating object have the same angular velocity, but do not have the same tangential velocity unless they are the same distance from the rotational axis. Tangential speed increases as the distance from the axis of rotation increases.

Tangential acceleration is the change in tangential speed per unit time. Its magnitude increases as the radius increases. Its direction is the same as the tangential velocity if the object is speeding up, and in the opposite direction as the velocity if it is slowing down.

You have now studied three types of acceleration relating to rotational motion. Centripetal acceleration is due to the change in direction of an object in circular motion. Tangential acceleration is the linear acceleration due to a change in angular velocity. Angular acceleration is the rate of change in angular velocity.

You can add the tangential and centripetal acceleration vectors of an object to determine the total linear acceleration of an object in rotational or circular motion.

Equations

$$\theta = \frac{s}{r}$$

$$\Delta\theta = \theta_f - \theta_i$$

$$\bar{\omega} = \frac{\Delta\theta}{\Delta t}$$

$$\omega = \lim_{\Delta t \to 0} \frac{\Delta\theta}{\Delta t}$$

$$\bar{\alpha} = \frac{\Delta\omega}{\Delta t}$$

$$\alpha = \lim_{\Delta t \to 0} \frac{\Delta\omega}{\Delta t}$$

$$\omega_f = \omega_i + \alpha t$$

$$\Delta\theta = \omega_i t + \frac{1}{2}\alpha t^2$$

$$\omega_f^2 = \omega_i^2 + 2\alpha\Delta\theta$$

$$\Delta\theta = \frac{1}{2}(\omega_f + \omega_i)t$$

$$v_T = r\omega$$

$$a_T = r\alpha$$

Chapter 10 Problems

Conceptual Problems

C.1 Is it possible for a rotating object to have increasing angular speed and negative angular acceleration? Explain your answer.

○ Yes ○ No

C.2 Order these three cities from smallest to largest tangential velocity due to the rotation of the Earth: Washington, DC, USA; Havana, Cuba; Ottawa, Canada.

smallest:
- i. Havana
- ii. Ottawa
- iii. Washington

middle:
- i. Havana
- ii. Ottawa
- iii. Washington

largest:
- i. Havana
- ii. Ottawa
- iii. Washington

C.3 Which of the following rotational quantities are the same for all points on a rotating disk? Check all that apply, and explain your selections.

- ☐ Angular velocity
- ☐ Tangential velocity
- ☐ Angular acceleration
- ☐ Tangential acceleration
- ☐ Centripetal acceleration

C.4 Does the angular velocity vector of the Earth point north or south along its axis of rotation?

○ North ○ South

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) If you increase the period, will the angular velocity increase, decrease or stay the same? (b) If you increase the period, will the linear speed increase, decrease or stay the same? (c) If you increase the distance from the center, will the angular velocity increase, decrease or stay the same? (d) If you increase the distance from the center, will the linear speed increase, decrease or stay the same?

(a)
- i. Increase
- ii. Stay the same
- iii. Decrease

(b)
- i. Increase
- ii. Stay the same
- iii. Decrease

(c)
- i. Increase
- ii. Stay the same
- iii. Decrease

(d)
- i. Increase
- ii. Stay the same
- iii. Decrease

0.2 Using the simulation in the interactive problem in this section and referring to your answers to the previous problem, what is the best way to maximize the linear speed of the rocket? Test your answer using the simulation.

- i. Maximize both the period and the distance from the center
- ii. Maximize the period and minimize the distance from the center
- iii. Minimize both the period and the distance from the center
- iv. Minimize the period and maximize the distance from the center

Section 1 - Angular position

1.1 Two cars are traveling around a circular track. The angle between them, from the center of the circle, is 55° and the track has a radius of 50 m. How far apart are the two cars, as measured around the curve of the track?

__________ m

1.2 Glenn starts his day by walking around a circular track with radius 48 m for 15 minutes. First he walks in a counterclockwise direction for 1000 meters, then he walks clockwise until the 15 minutes are up. This morning, his clockwise walk is 880 meters long. When he ends his walk, what is his angular position with respect to where he starts?

__________ rad

Section 2 - Angular displacement

2.1 A dancer completes 2.2 revolutions in a pirouette. What is her angular displacement?

__________ rad

2.2 What is the angular displacement in radians that the minute hand of a watch moves through from 3:15 A.M. to 7:30 P.M. the same day? Express your answer to the nearest whole radian.

__________ rad

2.3 What is the angular displacement in radians of the Earth around the Sun in one hour? Assume the orbit is circular and takes exactly 365 days in a counterclockwise direction, as viewed from above the North Pole.

__________ rad

2.4 The radius of the tires on your car is 0.33 m. You drive 1600 m in a straight line. What is the angular displacement of a point on the outer rim of a tire, around the center of the tire, during this trip? Assume the tire rotates in the counterclockwise direction.

__________ rad

2.5 A heavy vault door is shut. The angular position of the door from t = 0 to the time the door is shut is given by $\theta(t) = 0.125t^2$, where θ is in radians. (a) The door is completely shut at $\theta = \pi/2$ radians. At what time does this occur? (b) What is the angular displacement of the door between t = 0.52 s and t = 1.67 s? (c) What is the door's average angular velocity between t = 1.50 s and t = 2.50 s?

(a) __________ s
(b) __________ rad
(c) __________ rad/s

Section 3 - Angular velocity

3.1 A hamster runs in its wheel for 2.7 hours every night. If the wheel has a 6.8 cm radius and its average angular velocity is 3.0 radians per second, how far does the hamster run in one night?

__________ m

3.2 An LP record rotates at 33 1/3 rpm (revolutions per minute) and is 12.0 inches in diameter. What is the angular velocity in rad/s for a fly sitting on the outer edge of an LP rotating in a clockwise direction?

__________ rad/s

3.3 What is the average angular velocity of the Earth around the sun? Assume a circular counterclockwise orbit, and 365 days in a year.

__________ rad/s

3.4 A car starts a race on a circular track and completes the first three laps in a counterclockwise direction in 618 seconds, finishing with an angular velocity of 0.0103 rad/s. What is the car's average angular velocity for the first three laps?

__________ rad/s

3.5 Your bicycle tires have a radius of 0.33 m. It takes you 850 seconds to ride 14 times counterclockwise around a circular track of radius 73 m at constant speed. (a) What is the angular velocity of the bicycle around the track? (b) What is the magnitude of the angular velocity of a tire around its axis? (That is, don't worry about whether the tire's rotation is clockwise or counterclockwise.)

(a) __________ rad/s
(b) __________ rad/s

Section 4 - Angular acceleration

4.1 The blades of a fan rotate clockwise at −225 rad/s at medium speed, and −355 rad/s at high speed. If it takes 4.65 seconds to get from medium to high speed, what is the average angular acceleration of the fan blades during this time?

__________ rad/s^2

4.2 The blades of a kitchen blender rotate counterclockwise at 2.2×10^4 rpm (revolutions per minute) at top speed. It takes the blender 2.1 seconds to reach this top speed after being turned on. What is the average angular acceleration of the blades?

__________ rad/s^2

4.3 A potter's wheel is rotating at 3.2 rad/s in a counterclockwise direction when the potter turns it off and lets it slow to a stop. This takes 26 seconds. What is the average angular acceleration of the wheel during this time?

__________ rad/s^2

Section 7 - Equations for rotational motion with constant acceleration

7.1 A merry-go-round is at rest before a child pushes it so that it rotates with a constant angular acceleration for 27.0 s. When the child stops pushing, the merry-go-round is rotating at 1.20 rad/s. How many revolutions did the child make around the merry-go-round while he was pushing it?

__________ rev

7.2 A CD is rotating counterclockwise at 31 rad/s. What angular acceleration will bring it to a stop in 28 rad?

__________ rad/s^2

7.3 A motorcycle rider starts from rest and goes 3.75 rad counterclockwise around a circular track in 12.3 seconds. Assuming she accelerates at a constant rate, what is her angular acceleration?

__________ rad/s^2

7.4 A rotating water pump works by taking water in at one side of a rotating wheel, and expelling it from the other side. If a pump with a radius of 0.120 m starts from rest and accelerates at 30.5 rad/s^2, how fast will the water be traveling when it leaves the pump after it has been accelerating for 9.00 seconds?

__________ m/s

7.5 A cyclist starts from rest and rides in a straight line, increasing speed so that her wheels have a constant angular acceleration of 2.0 rad/s^2 around their axles. She accelerates until her wheels are rotating at 8.0 rad/s. If the radius of a tire is 0.29 meters, how far has the cyclist traveled?

__________ m

7.6 Many telescopes are housed in observatory domes, and a slit in the dome is opened to allow the telescope to see the sky. The dome has to be rotated so that the slit lines up with the telescope. In one observatory, there is a motor which causes the dome to rotate counterclockwise at 0.0800 rad/s. When the motor is shut off, the dome continues to rotate, but with an angular acceleration of −0.0400 rad/s^2 until the angular velocity is zero. The dome is currently rotating, and the telescope is pointed at an angle of positive 1.30 radians from north. At what angle from north should the slit be located when you shut off the motor, so that the slit lines up exactly with the telescope when the dome stops moving?

__________ rad

7.7 An athlete jogs at a constant angular velocity of 0.150 rad/s around a circular track. At time t = 0, she passes a stationary runner, who immediately starts chasing her with a constant angular acceleration of 0.225 rad/s^2. At what time will the second runner have caught up to the first runner?

__________ s

7.8 A coin is rolled without slipping on a table top in a straight line. It starts rolling at 3.4 rad/s and slows down at a constant angular acceleration. It is rolling at 1.2 rad/s when it falls off the table edge. If the radius of the coin is 0.011 m, and the edge of the table is 1.6 m from where the coin started, for how much time did the coin roll?

__________ s

7.9 The platter of a modern hard disk drive spins at 7.20×10^3 rpm (revolutions per minute). (a) How much time, in seconds, does it take for the disk to make a complete revolution? (b) Starting from rest, suppose the disk reaches full speed in 5.00 seconds. What is the average angular acceleration of the disk in radians per second? (c) Assuming constant angular acceleration, how many revolutions has the hard disk turned while spinning up to its final angular velocity?

(a) __________ s
(b) __________ rad/s^2
(c) __________ rev

7.10 Two cars race around a circular track. Car A accelerates at 0.340 rad/s^2 around the track, and car B at 0.270 rad/s^2. They start at the same place on the track and car A lets the slower-to-accelerate car B start first. Car B starts at time t = 0. When car A starts, car B has an angular velocity of 1.40 rad/s. At what time does car A catch up to car B?

__________ s

Section 10 - Interactive problem: launch the rocket

10.1 Use the information given in the interactive problem in this section to answer the following question. What is the angular acceleration required for the rocket to reach the desired angular velocity after one revolution? Test your answer using the simulation.

__________ rad/s^2

Section 11 - Tangential velocity

11.1 How might a magician make the Statue of Liberty disappear? Imagine that you are sitting with some spectators on a circular platform that, unknown to all of you, can rotate very slowly. It is evening, and you can see the Statue of Liberty a short distance away between two tall brightly lit columns at the rim of the platform. A large curtain can be drawn between the columns to temporarily hide the statue. The magician closes the curtain, then rotates the platform through an angle of just 0.170 radians so the statue is hidden behind one of the columns when the curtain is opened. (a) If the platform rotation takes 24.0 seconds, what is the average angular speed required? (b) You are sitting 4.00 m from the center of rotation while the platform is rotating. What is the centripetal acceleration required to move you along the circular arc? (c) Calculate the centripetal acceleration as a fraction of *g*. You could be unaware of the rotation, especially if you were distracted.

(a) __________ rad/s
(b) __________ m/s^2 (c) __________ g

11.2 An old-fashioned LP record rotates at 33 1/3 rpm (revolutions per minute) and is 12 inches in diameter. A "single" rotates at 45 rpm and is 7.0 inches in diameter. If a fly sits on the edge of an LP and then on the edge of a single, on which will the fly experience the greater tangential speed?

○ On the LP ○ On the 45

11.3 A computer hard drive disk with a diameter of 3.5 inches rotates at 7200 rpm. The "read head" is positioned exactly halfway from the axis of rotation to the outer edge of the disk. What is the tangential speed in m/s of a point on the disk under the read head?

__________ m/s

11.4 A radio-controlled toy car has a top speed of 7.90 m/s. You tether it to a pole with a rigid horizontal rod and let it drive in a circle at top speed. (a) If the rod is 1.80 m long, how long does it take the car to complete one revolution? (b) What is the angular velocity?

(a) __________ s
(b) __________ rad/s

11.5 You accelerate your car from rest at a constant rate down a straight road, and reach 22.0 m/s in 111 s. The tires on your car have radius 0.320 m. Assuming the tires rotate in a counterclockwise direction, what is the angular acceleration of the tires?

__________ rad/s^2

11.6 When a compact disk is played, the angular velocity varies so that the tangential speed of the area being read by the player is constant. If the angular velocity of the CD when the player is reading at a distance of 3.00 cm from the center is 3.51 revolutions per second, what is the angular velocity when the player is reading at a distance of 4.00 cm from the center?

__________ rev/s

Section 12 - Tangential acceleration

12.1 A whirling device is launched spinning counterclockwise at 35 rad/s. It slows down with a constant angular acceleration and stops after 16 seconds. If the radius of the device is 0.038 m, what is the magnitude of the tangential acceleration of a point on the edge of the device?

__________ m/s^2

12.2 Two go-karts race around a course that has concentric circular tracks. The radius of the inner track is 15.0 m, and the radius of the outer track is 19.0 m. The go-karts start from rest at the same angular position and time, and move at the same constant angular acceleration. The race ends in a tie after one complete lap, which takes 21.5 seconds. (a) What is the common angular acceleration of the carts? (b) What is the tangential acceleration of the inner cart? (c) What is the tangential acceleration of the outer cart?

(a) ________ rad/s^2
(b) ________ m/s^2
(c) ________ m/s^2

12.3 A car starts a race from rest on a circular track and has a tangential speed of 43 m/s at the end of the third lap. The track has a radius of 91 m. If it has constant angular acceleration, what is the magnitude of its tangential acceleration?

________ m/s^2

Section 14 - Tangential and centripetal acceleration

14.1 A race car drives around a circular track of radius 240 m at a constant tangential speed of 64 m/s. (a) What is the magnitude of the car's total acceleration? (b) What angle does the total acceleration vector make with the centripetal acceleration vector?

(a) ________ m/s^2
(b) ________ °

14.2 A windmill starts from rest and rotates with a constant angular acceleration of 0.25 rad/s^2. How many seconds after starting will the magnitude of the tangential acceleration of the tip of a blade equal the magnitude of the centripetal acceleration at the same point?

________ s

14.3 A car starts from rest and drives around a circular track with a radius of 45.0 m at constant tangential acceleration. If the car takes 27.0 s in its first lap around the track, (a) what is the magnitude of its overall acceleration at the end of that lap? (b) What angle does the acceleration vector make with a radius line?

(a) ________ m/s^2
(b) ________ °

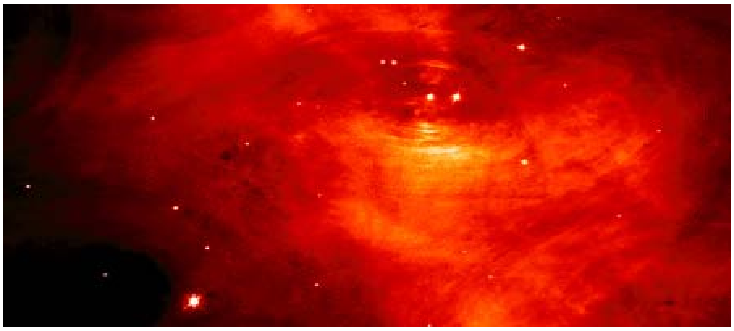
Section 16 - Interactive summary problem: 11.6 seconds to liftoff

16.1 Use the information given in the interactive problem in this section to calculate (a) the angular acceleration and (b) the distance of the rocket from the pivot that is required for the rocket to lift off in the stated time after exactly 2 revolutions. Express your answer to the nearest half meter. Test your answer with the simulation.

(a) ________ rad/s^2
(b) ________ m

Additional Problems

A.1 A pulsar is a rapidly rotating neutron star. The Crab Pulsar is located in the Crab Nebula in the constellation Taurus. The pulsar is in the center of the close-up view of the nebula shown in the photograph. The periods of pulsars can be measured with great accuracy: The Crab Pulsar has a period of 0.033 s. (a) Find the pulsar's angular velocity. (b) The radius of the pulsar is estimated to be 10 km. Find the tangential velocity of a point on its equator.



(a) ________ rad/s
(b) ________ m/s

A.2 You are designing an uninhabited combat air vehicle (UCAV) that will be capable of making a 20 "gee" turn. That is, the magnitude of the centripetal acceleration during the turn can be as great as 20.0 times 9.80 m/s^2. Assume that your UCAV flies at a speed of 331 m/s ("Mach 1") and that its mass is 5.00×10^3 kg. (a) What is the minimum radius of a horizontal turn that your UCAV can make? (b) What is the force ("thrust") in the horizontal direction that must be provided to make that turn?

(a) ________ m
(b) ________ N

A.3 A carnival game is set up so that two thin disks of equal size are fixed to the same horizontal axle, rotating at 2.5 rad/s. The disks are rotating above a frictionless table, and their rims just brush against the table. Each disk has a rectangular hole notched in it at the rim, and the holes have an angular separation of 0.25 rad as viewed down the length of the axle. The spacing between the disks, along the axis, is 0.11 m. In order to win the game, you must slide a puck along the table, through both holes. The holes are just big enough for the puck to get through, and if the puck hits a disk, you lose. What is the speed a puck must travel so that it slides through both holes?

__________ m/s

A.4 A diver performs a dive starting from a handstand off a 10.0-meter platform. What is her average angular velocity during the dive if she completes exactly 3 revolutions before she hits the water?

__________ rad/s

A.5 John and Joan walk in opposite directions around a circular path, starting from the same point. The path has a radius of 50.0 meters. John walks at 1.00 m/s, Joan at 1.25 m/s. How long will it take for them to meet?

__________ s

A.6 Prove that for an object in uniform circular motion, the centripetal acceleration equals the radius times the square of the angular velocity.

chapter 11 Rotational Dynamics

11.0 - Introduction

In the study of rotational kinematics, you analyze the motion of a rotating object by determining such properties as its angular displacement, angular velocity or angular acceleration. In this chapter, you explore the origins of rotational motion by studying *rotational dynamics*.

You study what causes an object to rotate more or less quickly, and how this relates to rotational work, rotational energy and angular momentum. This parallels a prior sequence, where you started with linear motion (kinematics), and then moved on to study force, work, energy and so on (dynamics). You will discover many similarities between linear and rotational dynamics, as well as some crucial differences.

At the right is a simulation that lets you conduct some experiments in the arena of rotational dynamics. In it, you play the role of King Kong, and your mission is to save the day, namely, the bananas on the truck. The bridge is initially open, and the truck loaded with bananas is heading toward it. You must rotate the bridge to a closed position. You determine where on the bridge you push and with how much force. If you cause the bridge to rotate too slowly, it will not close in time, and the truck will fall into the river. If you accelerate the bridge at too great a rate, the bridge will smash through the pilings.

interactive 1

Set Kong's force and his distance from the pivot to close the bridge in time

In this simulation, you are experimenting with torque, the rotational analog to force. A net force causes linear acceleration, and a net torque causes angular acceleration. The greater the torque you apply on the bridge, the greater the angular acceleration of the bridge. You control two of the elements that determine torque: the amount of force and how far it is applied from the axis of rotation. The third factor, the angle at which the force is applied, is a constant $90°$ in this simulation.

Try pushing with the same amount of force at different points on the bridge. Is the angular acceleration the same or different? Where do you push to create the maximum torque and angular acceleration? Select a combination of force and location that swings the bridge closed before the truck arrives, but not so hard that the pilings get smashed.

11.1 - Torque

Torque: A force that causes or opposes rotation.

A net force causes linear acceleration: a change in the linear velocity of an object. A net torque causes angular acceleration: a change in the angular velocity. For instance, if you push hard on a wrench like the one shown in Concept 1, you will start it and the nut rotating.

We will use a wrench that is loosening a nut as our setting to explain the concept of torque in more detail. In this section, we discuss two of the factors that determine the amount of torque. One factor is how much force F is exerted and the other is the distance r between the axis of rotation and the location where the force is applied. We assume in this section that the force is applied perpendicularly to the line from the axis of rotation and the location where the force is applied. (If this description seems cryptic, look at Concept 1, where the force is being applied in this manner.)

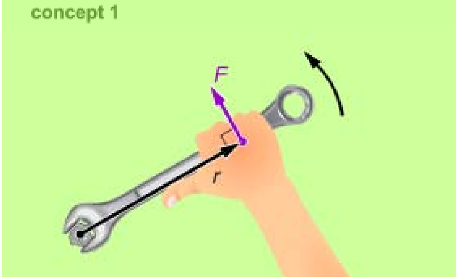
When the force is applied as stated above, the torque equals the product of the force F and the distance r. In Equation 1, we state this as an equation. The Greek letter τ (tau) represents torque.

Your practical experience should confirm that the torque increases with the amount of force and the distance from the axis of rotation. If you are trying to remove a “frozen” nut, you either push harder or you get a longer wrench so you can apply the force at a greater distance.

The location of a doorknob is another classic example of factoring in where force is applied. A torque is required to start a door rotating. The doorknob is placed far from the axis of rotation at the hinges so that the force applied to opening the door results in as much torque as possible. If you doubt this, try opening a door by pushing near its hinges.

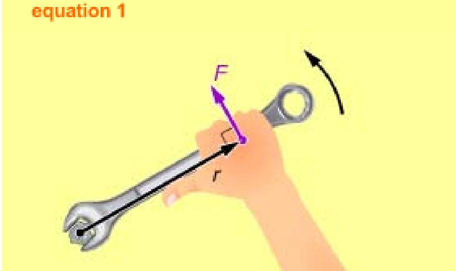
The wrench and nut scenario demonstrates another aspect of torque. The angular acceleration of the nut is due to a **net** torque. Let's say the nut in Concept 1 is stuck: the force of static friction between it and the bolt creates a torque that opposes the torque caused by the force of the hand. If the hand pushes hard enough and at a great enough distance from the nut, the torque it causes will exceed that caused by the force of static

concept 1



Torque
Causes or opposes rotation
Increases with:
· amount of force
· distance from axis to point of force

equation 1



For force applied perpendicularly

friction, and the nut will accelerate and begin rotating. The torque caused by the force of kinetic friction will continue to oppose the motion.

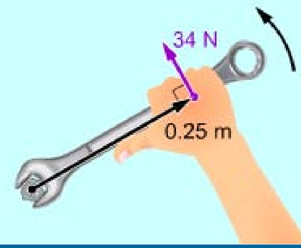
A net torque can cause an object to start rotating clockwise or counterclockwise. By convention, a torque that would cause counterclockwise rotation is a positive torque. A negative torque causes clockwise rotation. In Example 1, the torque caused by the hand on the wrench is positive, and the torque caused by friction between the nut and bolt is negative.

The unit for torque is the newton-meter (N·m). You might notice that work and energy are also measured using newton-meters, or, equivalently, joules. Work (and energy) and torque are different, however, and to emphasize that difference, the term "joule" is not used when discussing torque, but only when analyzing work or energy.

$\tau = rF$

τ = magnitude of torque
r = distance from axis to force
F = force
Counterclockwise +, clockwise –
Units: newton-meters (N·m)

example 1



The hand applies a force of 34 N as shown. Static friction creates an opposing torque of 8.5 N·m. Does the nut rotate?

$\tau = rF$

$\tau = (0.25\text{ m})(34\text{ N})$

$\tau = 8.5\text{ N·m}$

No, the nut does not rotate

11.2 - Torque, angle and lever arm

In the last section, we used a wrench to demonstrate torque. The force was applied perpendicularly to the wrench, but that need not be the case. Here we use a different example, a child sitting on a seesaw, to illustrate the non-perpendicular case. As shown in Equation 1, we treat both $\mathbf{F}$ and $\mathbf{r}$ as vectors, so we can describe the angle between them. The position vector $\mathbf{r}$ points from the axis of rotation to the point where the force is applied.

The child's weight creates a torque on the seesaw.

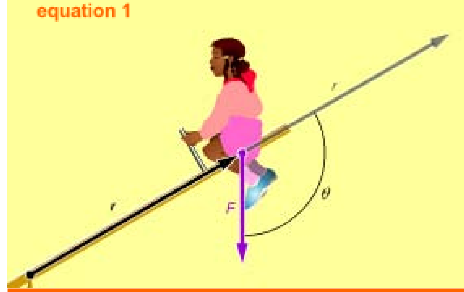
As before, $\mathbf{F}$ and $\mathbf{r}$ are two factors in determining the amount of torque applied. In this situation, $\mathbf{F}$ equals the child's weight, which points straight down. A third factor also determines how much torque she exerts: the sine of the angle θ between $\mathbf{r}$ and $\mathbf{F}$.

In Equation 1, we show the equation for calculating torque. Torque equals $\mathbf{r} \times \mathbf{F}$. The "×" stands for a vector operation called the cross product. The result of the cross product is another vector. The second equation shows how to calculate the magnitude of the cross product. The amount of torque equals the product of r, F and the sine of θ. To calculate torque, it is important to choose the correct angle for θ. It is the angle made by the two vectors when they are placed tail-to-tail. In Equation 1, you can see how to determine θ by moving $\mathbf{r}$ so it is tail-to-tail with $\mathbf{F}$. (The angle used is the **smaller** angle made between the two vectors, not the angle that would be made here by sweeping counterclockwise from $\mathbf{r}$ to $\mathbf{F}$.)

$F \sin\theta$ is the component of the force perpendicular to $\mathbf{r}$, so torque is the "perpendicular force" times the distance r. Any component that is parallel to the seesaw's plank does not cause it to rotate.

Torque is also frequently calculated in terms of a *lever arm*. You see this in Equation 2. The lever arm is the perpendicular distance from the axis of rotation to the *line of action*, which is the line containing the force vector. Since the lever arm equals the product $r \sin\theta$, the torque is the product of the force and the lever arm. In other words, the formula for torque using the lever arm is a rearrangement of the factors of the cross

equation 1



Torque

$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$

$\tau = rF \sin\theta$

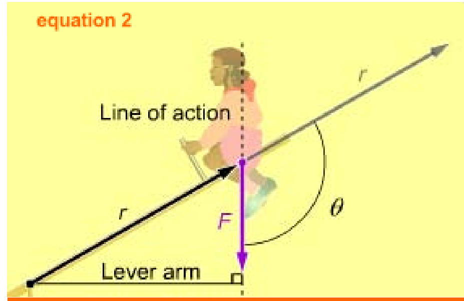
$\boldsymbol{\tau}$ = torque

product in Equation 1.

Children are sophisticated about torques, whether they know it or not. They understand that torques can be added. For example, if two children sit on the same side of a seesaw, their torques combine to create a larger net torque than that supplied by one child alone. If they sit on opposite sides, the net torque is less than either child's torque alone.

Children also learn that they can adjust the amount of torque they apply by moving toward or away from the axis of rotation. This means two children with different weights can balance each other, since both torques are a function of their weights and their distances from the axis of rotation. The heavier child slides closer to the axis, and the net torque is zero.

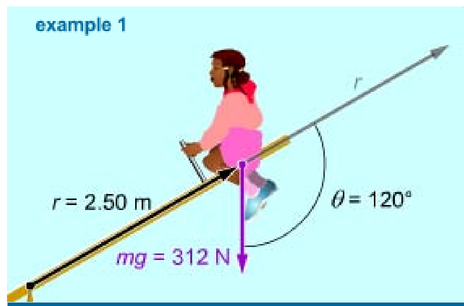
r = position vector
F = force
θ = angle between **r** and **F**
Units: newton-meters (N·m)



Calculating torque using lever arm

$\tau = F\,(r \sin \theta)$

$r \sin \theta$ = lever arm



What is the torque exerted by the girl?

$\tau = rF \sin \theta$

τ = (2.50 m)(312 N)(sin 120°)

τ = –675 N·m (clockwise)

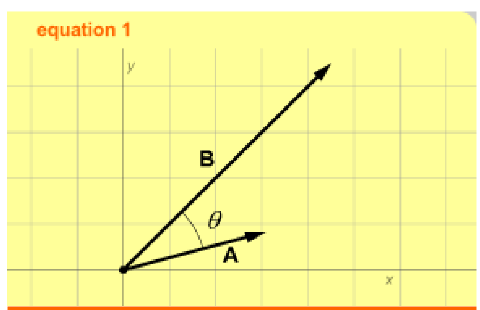
11.3 - Cross product of vectors

Cross product: A vector whose magnitude equals the product of the magnitudes of two vectors and the sine of the smaller angle between them. Its direction is determined by the right-hand rule.

Several physics properties, including torque, are calculated using the cross product. The cross product is a way to multiply two vectors. The result is a vector that is sometimes called their *vector product*. To determine the magnitude of the cross product, multiply the product of the magnitudes of the two vectors by the sine of the angle between them. This formula is shown on the right. As the diagram shows, placing vectors tail-to-tail will allow you to determine the correct angle. The angle used is the smaller angle between the two vectors.

A technique called the *right-hand rule* will help you determine the direction of the vector that results from the cross product. (Right-hand rules are also frequently used in the study of electricity and magnetism.)

How to apply the right-hand rule is shown on the right (you and your classmates may



Magnitude of A × B

$AB \sin \theta$

A = magnitude of vector **A**
B = magnitude of vector **B**
θ = smaller angle between **A** and **B**

soon find yourselves making these awkward gestures). First, make sure the vectors are placed tail-to-tail. Then, point the fingers of your right hand along the first vector, and curl them toward the second vector. (Curl your fingers through the smaller of the two angles between the vectors.) The direction of your thumb is the direction of the resulting vector: It is perpendicular to both of the two vectors you are multiplying.

Often, the vectors are depicted as being on the surface of a book or computer monitor. This means the cross product will point directly toward you out of the page or monitor, or it will point away from you.

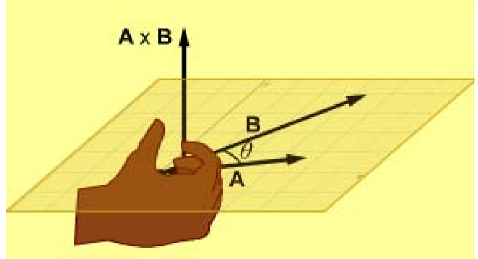
This same point can be stated mathematically: Both of the vectors being multiplied are typically considered to lie in the xy plane of a graph, and to start at the origin. For instance, one of the vectors might stretch from the origin to the point (2, 3). The resulting cross product vector points toward or away from you along the z axis.

As mentioned, a positive torque causes a stationary object to rotate in a counterclockwise fashion, and a negative torque does the opposite. Torques must be added as vectors. A torque acting on an object with a fixed axis of rotation has only two possible vector directions. Using positive and negative signs is a convenient way to do this vector addition. Adding two vectors pointing in the same direction results in a net vector with a greater magnitude.

We will mention a few other mathematical properties of cross products. Since the order of the vectors dictates the direction determined by the right-hand rule, the cross product is not commutative. In fact, $\mathbf{A} \times \mathbf{B} = -(\mathbf{B} \times \mathbf{A})$. The vectors that result from these cross products have the same magnitude, but they point in opposite directions.

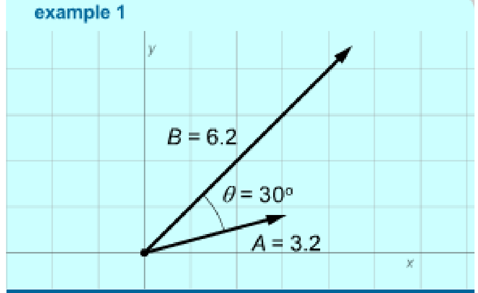
There are two other useful facts about cross products. First, $\mathbf{A} \times \mathbf{A} = \mathbf{0}$, the zero vector. Why? The angle in this case is zero, and $\sin(0) = 0$. $\mathbf{A} \times -\mathbf{B} = -\mathbf{A} \times \mathbf{B}$ is another useful identity.

equation 2



Direction of A × B
Use right-hand rule
· Place vectors tail-to-tail
· Curl fingers from **A** to **B**
· Thumb indicates direction

example 1



What is the cross product B × A?
Magnitude = $BA \sin\theta$
Magnitude = (6.2)(3.2)(sin 30°)
Magnitude = 9.9
Direction = away from you

11.4 - Torque, moment of inertia and angular acceleration

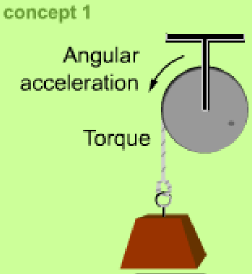
Moment of inertia: The measure of resistance to angular acceleration.

An object's moment of inertia is the measure of its resistance to a change in its angular velocity. It is analogous to mass for linear motion; a more massive object requires more net force to accelerate at a given rate than a less massive object. Similarly, an object with a greater moment of inertia requires more net torque to angularly accelerate at a given rate than an object with a lesser moment of inertia. For example, it takes more torque to accelerate a Ferris wheel than it does a bicycle wheel, for the same rate of acceleration.

To state this as an equation: The net torque equals the moment of inertia times the angular acceleration. This equation, $\Sigma\tau = I\alpha$, resembles Newton's second law, $\Sigma F = ma$. We sometimes refer to this equation as *Newton's second law for rotation*. The moment of inertia is measured in kilogram·meters squared (kg·m^2). Like mass, the moment of inertia is always a positive quantity.

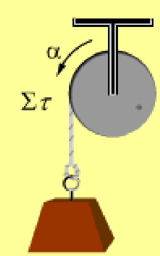
We show how the moment of inertia of an object could be experimentally determined in Example 1. A block, attached to a massless rope, is causing a pulley to accelerate. The angular acceleration and the net torque are stated in the problem. (The net torque could be determined by multiplying the tension by the radius of the pulley, keeping in mind that the tension is less than the weight of the block since the block accelerates downward.) With these facts known, the moment of inertia of the pulley can be determined.

concept 1



Torque and moment of inertia
Net torque = moment of inertia × angular acceleration

equation 1

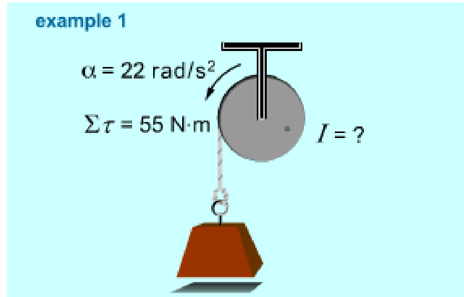


Torque and moment of inertia

$$\Sigma\tau = I\alpha$$

$\Sigma\tau$ = net torque
I = moment of inertia
α = angular acceleration
Units for I: kg·m^2

example 1



What is the moment of inertia of the pulley?

$\Sigma\tau = I\alpha$
$I = \Sigma\tau/\alpha$
I = (55 N·m)/(22 rad/s^2)
I = 2.5 kg·m^2

11.5 - Calculating the moment of inertia

If you were asked whether the same amount of torque would cause a greater angular acceleration with a Ferris wheel or a bicycle wheel, you would likely answer: the bicycle wheel. The greater mass of the Ferris wheel means it has a greater moment of inertia. It accelerates less with a given torque.

But more than the amount of mass is required to determine the moment of inertia; the distribution of the mass also matters. Consider the case of a boy sitting on a seesaw. When he sits close to the axis of rotation, it takes a certain amount of torque to cause him to have a given rate of angular acceleration. When he sits farther away, it takes more torque to create the same rate of acceleration. Even though the boy's (and the seesaw's) mass stays constant, he can increase the system's moment of inertia by sitting farther away from the axis.

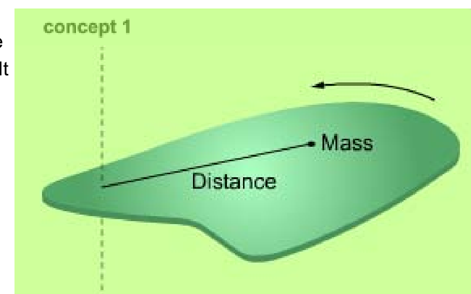
When a rigid object or system of particles rotates about a fixed axis, each particle in the object contributes to its moment of inertia. The formula in Equation 1 to the right shows how to calculate the moment of inertia. The moment equals the sum of each particle's mass times the square of its distance from the axis of rotation.

A single object often has a different moment of inertia when its axis of rotation changes. For instance, if you rotate a baton around its center, it has a smaller moment of inertia than if you rotate it around one of its ends. The baton is harder to accelerate when rotated around an end. Why is this the case? When the baton rotates around an end, more of its mass on average is farther away from the axis of rotation than when it rotates around its center.

If the mass of a system is concentrated at a few points, we can calculate its moment of inertia using multiplication and addition. You see this in Example 1, where the mass of the object is concentrated in two balls at the ends of the rod. The moment of inertia of the rod is very small compared to that of the balls, and we do not include it in our calculations. We also consider each ball to be concentrated at its own center of mass when measuring its distance from the axis of rotation (marked by the ×). This is a reasonable approximation when the size of an object is small relative to its distance from the axis.

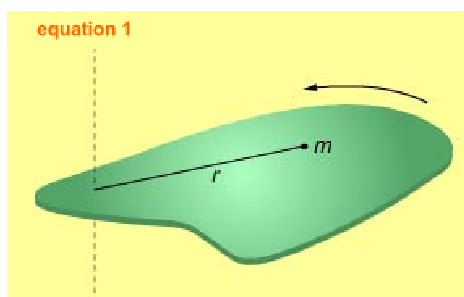
Not all situations lend themselves to such simplifications. For instance, let's assume we want to calculate the moment of inertia of a CD spinning about its center. In this case the mass is uniformly distributed across the entire CD. In such a case, we need to use calculus to sum up the contribution that each particle of mass makes to the moment, or we must take advantage of a table that tells us the moment of inertia for a disk rotating

concept 1



Moment of inertia
Sum of each particle's
· Mass times its
· Distance squared from the axis

equation 1



Moment of inertia

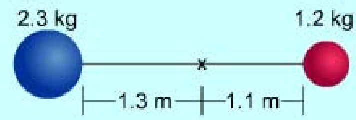
$$I = \Sigma mr^2$$

I = moment of inertia

around its center.

m = mass of a particle
r = distance of particle from axis
Units: kg·m²

example 1



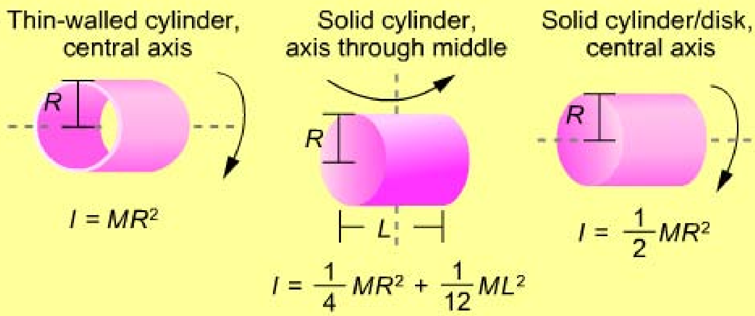
What is the system's moment of inertia? Ignore the rod's mass.

$I = \Sigma mr^2 = m_1r_1^2 + m_2r_2^2$

$I = (2.3\text{ kg})(1.3\text{ m})^2 + (1.2\text{ kg})(1.1\text{ m})^2$

$I = 5.3\text{ kg·m}^2$

11.6 - A table of moments of inertia



equation 1

Moments of inertia
Cylinders

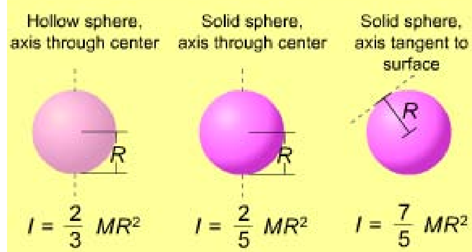
Sets of objects are shown in the illustrations above and to the right. Above each object is a description of it and its axis of rotation. Below each object is a formula for calculating its moment of inertia, I.

The variable M represents the object's mass. It is assumed that the mass is distributed uniformly throughout each object.

If you look at the formulas in each table, they will confirm an important principle underlying moments of inertia: The distribution of the mass relative to the axis of rotation matters. For instance, consider the equations for the hollow and solid spheres, each of which is rotating about an axis through its center. A hollow sphere with the same mass and radius as a solid sphere has a greater moment of inertia. Why? Because the mass of the hollow sphere is on average farther from its axis of rotation than that of the solid sphere.

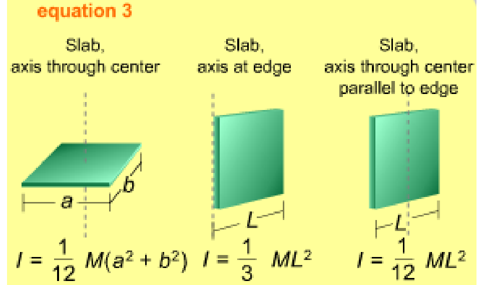
Note also that the moment of inertia for an object depends on the location of the axis of rotation. The same object will have different moments of inertia when rotated around differing axes. As shown on the right, a thin rod rotated around its center has one-fourth the moment of inertia as the same rod rotated around one end. Again, the difference is due to the distribution of mass relative to the axis of rotation. On average, the mass of the rod is further away from the axis when it is rotated around one end.

equation 2

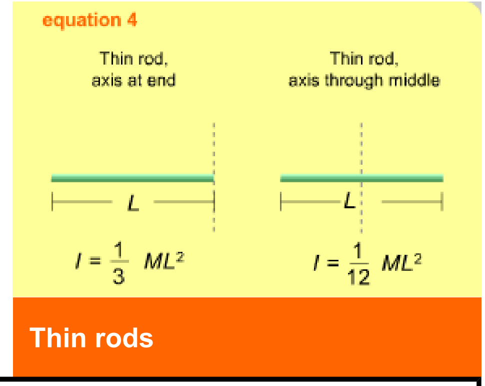


Spheres

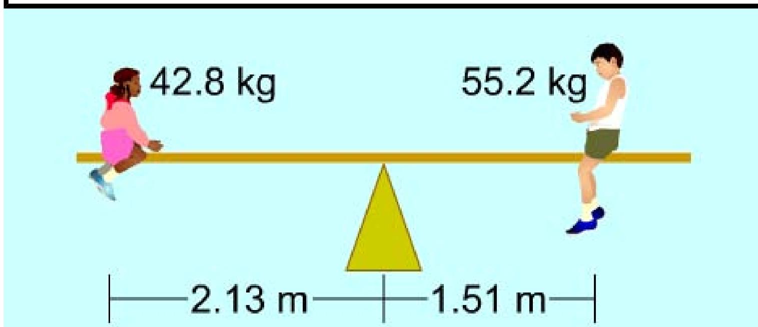
equation 3



Slabs



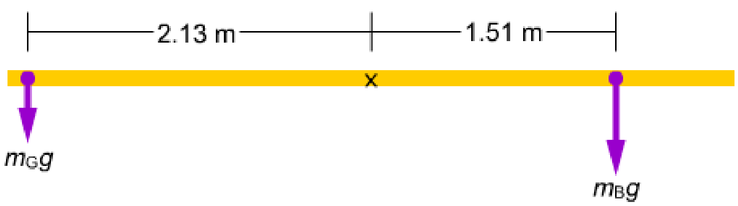
11.7 - Sample problem: a seesaw



The seesaw plank is horizontal. Its mass is 36.5 kg, and it is 4.40 m long. What is the initial angular acceleration of this system?

The axis of rotation is the point where the fulcrum touches the midpoint of the plank. The plank itself creates no net torque since it is balanced at its middle. For every particle at a given distance from the axis that creates a clockwise torque, there is a matching particle at the same distance creating a counterclockwise torque. However, the plank does factor into the moment of inertia.

Draw a diagram



Variables

mass of seesaw plank	$m_S = 36.5$ kg
seesaw plank's moment of inertia	I_S

	girl	boy
mass	$m_G = 42.8$ kg	$m_B = 55.2$ kg
distance from axis	$r_G = 2.13$ m	$r_B = 1.51$ m
moment of inertia	I_G	I_B

What is the strategy?

1. Calculate the moment of inertia of the system: the sum of the moments for the children, and the moment of the plank.
2. Calculate the net torque by summing the torques created by each child. The torques of the left and right sides of the plank cancel, so you do not have to consider them.
3. Divide the net torque by the moment of inertia to determine the initial angular acceleration.

Physics principles and equations

We will use the definitions of torque and moment of inertia.

$$\tau = rF \sin \theta, I = \Sigma mr^2$$

To calculate the moments of inertia of the children, we consider the mass of each to be concentrated at one point.

The plank can be considered as a slab rotating on an axis parallel to an edge through the center, with moment of inertia

$$I = \frac{1}{12} ML^2$$

The equation relating net torque and moment of inertia is Newton's second law for rotation,

$$\Sigma\tau = I\alpha$$

Step-by-step solution

First, we add the moments of inertia for the two children and the seesaw plank. The sum of these values equals the system's moment of inertia.

Step	Reason
1. $I = I_G + I_B + I_S$	total moment is sum
2. $I = m_G r_G^2 + m_B r_B^2 + I_S$	definition of moment of inertia
3. $I = m_G r_G^2 + m_B r_B^2 + \frac{1}{12} m_S L^2$	moment of slab
4. $I = (42.8)(2.13)^2 + (55.2)(1.51)^2 + \frac{1}{12}(36.5)(4.40)^2$	enter values
5. $I = 379\ \text{kg·m}^2$	evaluate

The children create torques, and to calculate the net torque, we sum their torques, being careful about signs. The plank creates no net torque since its midpoint is at the fulcrum.

Step	Reason
6. $\Sigma\tau = \tau_G + \tau_B$	net torque equals sum of torques
7. $\Sigma\tau = m_G g r_G + (-m_B g r_B)$	equation for torque
8. $\Sigma\tau = (42.8)(9.80)(2.13) - (55.2)(9.80)(1.51)$	enter values
9. $\Sigma\tau = 76.6\ \text{N·m}$	evaluate

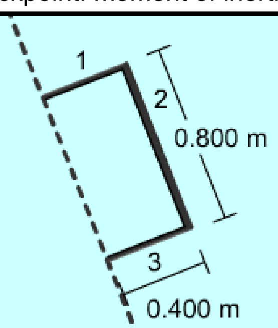
Now we use the values we calculated for the net torque and the moment of inertia to calculate the angular acceleration.

We divide to calculate the angular acceleration. Since the value is positive, the direction of rotation is counterclockwise.

Step	Reason
10. $\alpha = \Sigma\tau / I$	Newton's second law for rotation
11. $\alpha = (76.6\ \text{N·m})/(379\ \text{kg·m}^2)$	substitute equations 5 and 9 into equation 10
12. $\alpha = 0.202\ \text{rad/s}^2$ (counterclockwise)	solve for α

Because various quantities change, such as the angle between the direction of each child's weight and the seesaw, the angular acceleration changes as the seesaw rotates. This is why we asked for the **initial** angular acceleration.

11.8 - Interactive checkpoint: moment of inertia

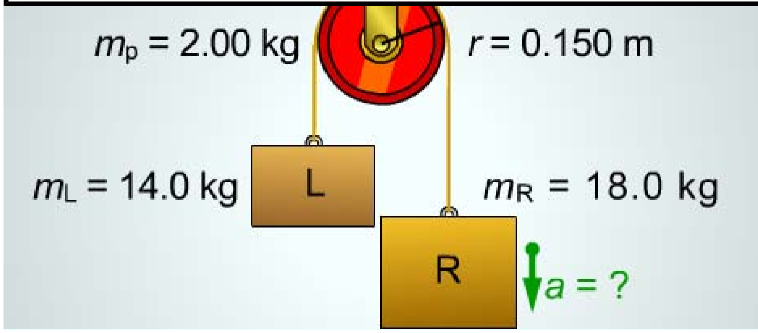


Three thin rods are connected in a U-shape and rotate about an axis as shown. Two rods are 0.400 m long and have a mass of 1.60 kg; the third rod is 0.800 m long with a mass of 3.20 kg. Find the moment of inertia of the U-shaped object as it rotates about this axis.

Answer:

I = ______ kg·m²

11.9 - Sample problem: an Atwood machine



What is the magnitude of the acceleration of the block on the right?

Here we consider an Atwood machine, factoring in the moment of inertia of the pulley. We will model the pulley as a uniform solid disk. We still assume that the rope is massless and does not stretch and that the pulley is frictionless.

The blocks' accelerations are equal and opposite, but the tension exerted on each block by the rope is different because of the pulley's moment of inertia. We will use the convention that downward acceleration and force are negative, and upward acceleration and force are positive. We know the block on the right will accelerate downward because it is more massive than the block on the left.

Variables

	Block L	Block R
tension	T_L	T_R
mass	$m_L = 14.0$ kg	$m_R = 18.0$ kg
weight	$-m_L g$	$-m_R g$
acceleration	a	$-a$

	Pulley
torque from left block	τ_L
torque from right block	τ_R
mass	$m_P = 2.00$ kg
radius	$r = 0.150$ m
moment of inertia	I
angular acceleration	α

What is the strategy?

1. Calculate the net force on each block. The forces are strictly vertical. Use Newton's second law for each block to find expressions for the tension force in the rope on each side of the pulley.
2. Calculate the net torque acting on the pulley. The forces that create torques on the pulley are the tensions of the rope. The tension forces are perpendicular to the radius of the pulley at the points where they act. Use Newton's second law for rotation to set $I\alpha$ equal to this net torque.
3. Use the definition of angular acceleration to introduce the linear acceleration of the masses into the previously derived equation, and solve for the linear acceleration.
4. Use the expression for the moment of inertia of a solid disk as the moment of inertia of the pulley, and use the values given in the problem to compute the acceleration.

Physics principles and equations

Newton's second law

$$\Sigma F = ma$$

Equations for torque

$$\tau = rF \sin \theta$$

$$\Sigma\tau = I\alpha$$

Moment of inertia of a solid disk

$$I = \tfrac{1}{2}mr^2$$

Angular acceleration

$$\alpha = a/r$$

The acceleration a is the positive magnitude of the acceleration of the blocks. Since the right block falls, this acceleration results in a clockwise (negative) angular acceleration of the pulley.

Step-by-step solution

We start by finding expressions for each of the tensions. The tension forces on each side of the pulley are not equal because of the pulley's moment of inertia and the unequal masses of the blocks.

Step	Reason
1. $\Sigma F_L = m_L a$	Newton's second law
2. $T_L + (-m_L g) = m_L a$	net force
3. $T_L = m_L g + m_L a$	solve for T_L
4. $\Sigma F_R = m_R a$	Newton's second law
5. $T_R + (-m_R g) = m_R(-a)$	net force
6. $T_R = m_R g - m_R a$	solve for T_R

Next, we find an expression for the net torque on the pulley and use it in Newton's second law for rotation. We ignore the upward force on the pulley from its support, because this force acts at the axis of rotation and creates no torque.

Step	Reason
7. $\Sigma\tau = \tau_L + \tau_R$	net torque
8. $\Sigma\tau = rT_L + (-rT_R)$	torque for perpendicular force
9. $\Sigma\tau = I\alpha$	Newton's second law for rotation
10. $I\alpha = rT_L - rT_R$	substitute from 9 into 8

Now we bring the pieces together. We substitute in the expressions for the tensions found above, as well as expressions for the angular acceleration and the moment of inertia. We are left with an equation for the acceleration in terms of the masses of the blocks and the pulley.

Step	Reason
11. $I\alpha = r(m_L g + m_L a) - r(m_R g - m_R a)$	substitute equations 3 and 6 into equation 10
12. $I\alpha = gr(m_L - m_R) + ar(m_L + m_R)$	rearrange right side
13. $\alpha = -a/r$	angular acceleration at edge of pulley (clockwise)
14. $-Ia/r = gr(m_L - m_R) + ar(m_L + m_R)$	substitute equation 13 into equation 12
15. $a = \dfrac{gr(m_R - m_L)}{I/r + rm_R + rm_L}$	solve for a

We use the formula for the moment of inertia for a solid disk, and substitute the known values to calculate the acceleration.

Step	Reason
16. $I = ½m_P r^2$	moment of inertia of pulley
17. $a = \dfrac{gr(m_R - m_L)}{½m_P r^2/r + rm_R + rm_L}$	Substitute equation 16 into equation 15
18. $a = \dfrac{2g(m_R - m_L)}{m_P + 2m_R + 2m_L}$	simplify
19. $a = \dfrac{2(9.80\ \text{m/s}^2)(18.0 - 14.0\ \text{kg})}{2.00\ \text{kg} + 2(18.0)\ \text{kg} + 2(14.0)\ \text{kg}}$	substitute values
20. $a = 1.19\ \text{m/s}^2$	evaluate

We find that the blocks accelerate at 1.19 m/s^2. In contrast, if the pulley's moment of inertia were zero, the blocks would accelerate at 1.23 m/s^2. You can confirm this by setting m_P equal to zero in step 19 above.

11.10 - Interactive problem: close the bridge

Once again, you are King Kong, and your task is to close the bridge you see on the right in order to save an invaluable load of bananas (well, invaluable to you at least). Here, we ask you to be a more precise gorilla than you may have been in the introductory exercise.

To close the bridge quickly enough to save the fruit without breaking off the bumper pilings, you need to apply a torque so that the bridge's angular acceleration is $\pi/16.0$ rad/s^2. The moment of inertia of the rotating part of the bridge is 45,400,000 kg·m^2.

Two trucks are parked on this part of the bridge, and you must include them when you calculate the total moment of inertia. Each truck has a mass of 4160 kg; the midpoint of one is 20.0 m and the midpoint of the other is 30.0 m from the pivot (axis of rotation) of the swinging bridge. The trucks will increase the bridge's moment of inertia. To solve the problem, consider all the mass of each truck to be concentrated at its midpoint.

You apply your force 35.0 m from the pivot and your force is perpendicular to the rotating component of the bridge. Enter the amount of force you wish to apply to the nearest 0.01×10^5 N and press GO to start the simulation. Press RESET if you need to try again.

If you have difficulty solving this problem, review the sections on calculating the moment of inertia, and the relation between torque, angular acceleration, and moment of inertia.

interactive 1

11.11 - Parallel axis theorem

The parallel axis theorem is a tool for calculating the moment of inertia of an object. You use it when you know the moment of inertia for an object rotating about an axis that passes through its center of mass and want to know the moment when it rotates around a different but parallel axis of rotation.

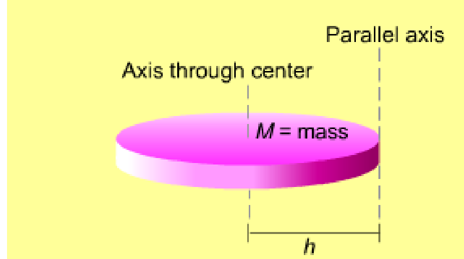
The illustration for Equation 1 shows two such parallel axes. The axis on the left passes through the center of mass of a cylindrical disk, the other is at the edge of the disk.

The theorem states that the moment of inertia when the disk rotates about the axis on its edge will be the sum of two values: the moment of inertia when the disk rotates about its center of mass, and the product of the disk's mass and the square of the distance between the two axes (shown as h in our diagram). This is stated as an equation to the right.

The usefulness of the parallel axis theorem lies in this fact: It is usually much easier to calculate the moment of inertia of an object around an axis through its center of mass than around an off-center axis. For example, if we must use an integral to calculate the moment of inertia, doing so around the center of mass lets us more readily take advantage of any symmetry of the object. The parallel axis theorem can then be used to find the moment of inertia around another parallel axis. In sum, the parallel axis theorem lets us use an easier integral and some algebra to calculate the moment for the parallel axis.

The disk to the right has a moment of inertia of $\frac{1}{2}MR^2$ when it rotates about its center. We can use this formula as the starting point in our calculation of the disk's moment of inertia when it is rotated about an axis at its edge. You see this computation worked out in Example 1.

equation 1

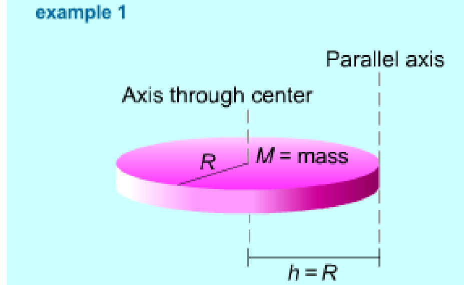


Parallel axis theorem

$$I_P = I_{CM} + Mh^2$$

I_P = moment of inertia, parallel axis
I_{CM} = moment, center of mass axis
M = mass
h = distance between axes

example 1



What is the moment of inertia of a disk when it rotates about the parallel axis shown above?

$I_P = I_{CM} + Mh^2$
$I_P = \frac{1}{2}MR^2 + MR^2$

$I_P = \frac{3}{2}MR^2$

11.12 - Rotational work

When you push or pull an object (apply a force) the work you do equals the product of the displacement and the force in the direction of the displacement.

With rotational motion, the work equals the product of the angular displacement and the torque. The two equations are analogous, and we start the derivation of the equation for rotational work with the definition of linear work.

In the derivation below, we use the scenario illustrated in Equation 2: A worker is pulling on a rope attached to the edge of a spool with force F. The spool rotates through an angular displacement $\Delta\theta$ but does not move linearly. The force is perpendicular to the radius from the axis of rotation to the point where the force is applied. The rotation of the spool does not materially affect the radius r.

Variables

rotational work done on spool	W
force applied to rope	F
amount of rope pulled out	Δx
radius of spool	r
angular displacement of spool	$\Delta\theta$
torque applied to spool	τ

Strategy

1. State the linear relationship between work, force and displacement.
2. Replace the force and displacement by their angular equivalents.
3. Simplify to get the rotational form of the work equation.

Physics principles and equations

The equation for linear work when the force is parallel to the displacement

$$W = F\Delta x$$

For a force perpendicular to the line from the axis of rotation

$$\tau = rF$$

Mathematics principle

The amount of rope pulled out equals the arc length labeled Δx on the illustration in Equation 2. This arc length is the radius of the spool times its angular displacement.

$$\Delta x = r\Delta\theta$$

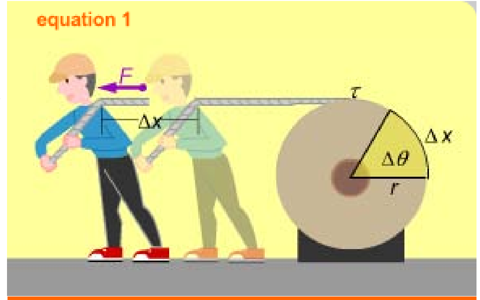
Step-by-step derivation

We start with the definitions of work, torque and the relationship of arcs and angles.

Step	Reason
1. $W = F\Delta x$	work
2. $F = \tau/r$	tangential force
3. $\Delta x = r\Delta\theta$	arc length and angle

In the following steps we substitute the rotational equivalents stated above into the work equation and simplify to obtain the desired result.

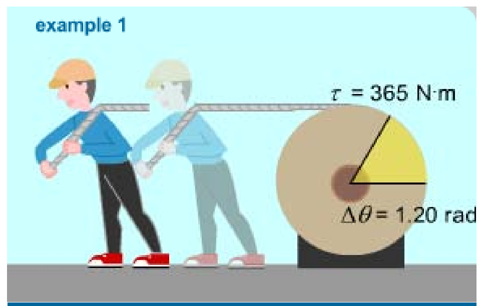
Step	Reason
4. $W = (\tau/r)(\ r\Delta\theta)$	substitute equations 2 and 3 into equation 1
5. $W = \tau\Delta\theta$	simplify



Rotational work

$$\boxed{W = \tau\Delta\theta}$$

W = work
τ = torque
$\Delta\theta$ = angular displacement



How much work is done when the wheel rotates?

$W = \tau\Delta\theta$

$W = (365\text{ N·m})(1.20\text{ rad})$

$W = 438\text{ J}$

11.13 - Rotational kinetic energy

The equation on the right enables you to calculate the rotational kinetic energy (KE) of a rigid, rotating object, the kinetic energy of an object due to its rotational motion. It is analogous to the equation for linear kinetic energy. The rotational KE equals one-half the moment of inertia times the square of the angular velocity.

This equation can be derived from the definition of linear kinetic energy. The rotating object consists of a large number of individual particles, each moving at a different linear (tangential) velocity. The kinetic energies of all the particles can be added to determine the kinetic energy of the entire object. To derive the equation to the right, the key insight is to see that the distance from the axis of rotation figures both in a particle's tangential velocity and in calculating its contribution to the disk's moment of inertia.

In the derivation, we start with a particle of mass m situated somewhere in a rigid object, as shown in the second illustration to the right. We derive the equation by first calculating the kinetic energy of the single particle. The total KE is the sum of the kinetic energies of all the particles.

Variables

mass of a particle on rotating object	m
linear speed of particle	v
distance of particle from axis of rotation	r
angular velocity of rotating object	ω
kinetic energy of particle	KE_p
kinetic energy of object	$KE = \Sigma KE_p$
moment of inertia of particle	I_p
moment of inertia of object	$I = \Sigma I_p$

Strategy

1. State the equation for the linear KE of a particle in terms of its speed. Restate the equation in rotational terms and simplify.
2. Rewrite the expression for the KE of a particle in terms of its moment of inertia.
3. Sum the individual kinetic energies of all the particles to derive the desired equation.

Physics principles and equations

The equation for the kinetic energy of a particle

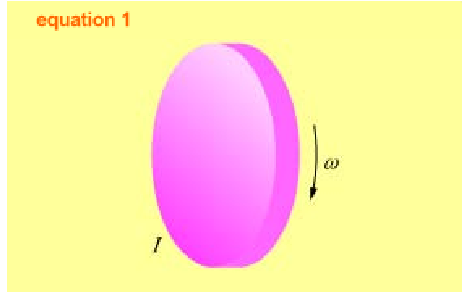
$$KE = ½mv^2$$

The relationship between linear speed and angular velocity is

$$v = \omega r$$

The moment of inertia of a single particle of mass m at distance r from the axis of rotation

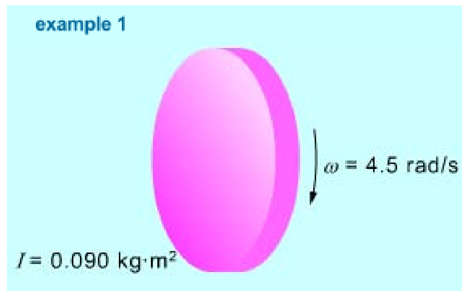
$$I = mr^2$$



Rotational kinetic energy

$$KE = ½I\omega^2$$

KE = rotational kinetic energy
I = moment of inertia
ω = angular velocity



What is the kinetic energy of the disk?

$KE = ½I\omega^2$
$KE = ½(0.090\ \text{kg·m}^2)(4.5\ \text{rad/s})^2$
$KE = 0.91$ J

Step-by-step derivation

For a particle on a rotating object, its linear speed as it moves along a circular path is its tangential speed. We use the definition of kinetic energy and the relation between linear and angular speed to write the kinetic energy of a particle in rotational terms.

Step	Reason
1. $KE_p = ½mv^2$	definition of kinetic energy
2. $v = \omega r$	linear speed and angular velocity
3. $KE_p = ½m(\omega^2 r^2)$	substitute equation 2 into equation 1
4. $KE_p = ½(mr^2)\omega^2$	rearrange

We next rewrite the particle's kinetic energy in terms of its moment of inertia.

Step	Reason
5. $I_p = mr^2$	moment of inertia
6. $KE_p = ½I_p\omega^2$	substitute equation 5 into equation 4

Finally, we sum the kinetic energy over all particles. This results in the desired equation for rotational KE.

Step	Reason
7. $KE = \Sigma KE_p$	total kinetic energy of object
8. $KE = \Sigma ½I_p\omega^2$	substitute equation 6 into equation 7
9. $KE = ½\omega^2\Sigma I_p$	factor
10. $KE = ½I\omega^2$	add moments and rearrange

In Example 1, you see an application of this equation.

11.14 - Physics at work: flywheels

Flywheels are rotating objects used to store energy as rotational kinetic energy. Recently, environmental and other concerns have caused flywheels to receive increased attention. Many of these new flywheels serve as mechanical “batteries,” replacing traditional electric batteries.

This advanced flywheel is being developed by NASA as a source of stored energy for use by satellites and spacecraft.

Why the interest? Traditional chemical batteries, rechargeable or not, have a shorter total life span than flywheels and can cause environmental problems when disposed of incorrectly. On the other hand, flywheels cost more to produce than traditional batteries, and their ability to function in demanding situations is unproven.

Flywheels can be powered by “waste” energy. For instance, when a bus slows down, its brakes warm up. The bus’s kinetic energy becomes thermal energy, which the vehicle cannot re-use efficiently. Some buses now use a flywheel to convert a portion of that linear kinetic energy into rotational energy, and then later transform that rotational energy back into linear kinetic energy as the bus speeds up.

Flywheels can receive power from more traditional sources, as well. For instance, uninterruptible power sources (UPS) for computers use rechargeable batteries to keep computers powered during short-term power outages. Flywheels are being considered as an alternative to chemical batteries in these systems. Less traditional sources can also supply energy to a flywheel: NASA uses solar power to energize flywheels in space.

The equation for rotational KE is shown to the right. The moment of inertia and maximum angular velocity determine how much energy a flywheel can store. The moment of inertia, in turn, is a function of the mass and its distance (squared) from the axis of rotation.

In Concept 1, you see a traditional flywheel. It is large, massive, and constructed with most of its mass at the outer rim, giving it a large moment of inertia and allowing it to store large amounts of rotational KE. In Concept 2, you see a modern flywheel, which is much smaller and less massive, but capable of rotating with a far greater angular velocity. Flywheels in these systems can rotate at 60,000 revolutions per minute (6238 rad/s). Air drag and friction losses are greatly reduced by enclosing the flywheel in a near vacuum and by employing magnetic bearings.

concept 1

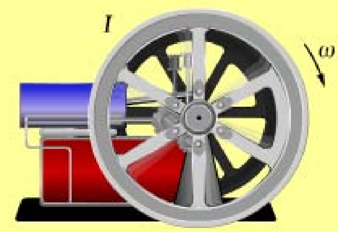
Flywheels
Spinning objects "store" rotational *KE*
Energy depends on
· angular velocity
· moment of inertia (mass, radius)

concept 2

Flywheels

Serve as mechanical batteries

equation 1



Flywheel energy

$$KE = ½I\omega^2$$

KE = kinetic energy (rotational)
I = moment of inertia
ω = angular velocity

11.15 - Rolling objects and kinetic energy

The coin shown in Equation 1 rolls without slipping. That is, it rotates and moves linearly as it travels to the right. Its total kinetic energy is the sum of its linear and rotational kinetic energies.

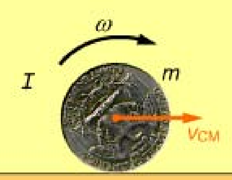
We can use two equations discussed earlier to determine the coin's total kinetic energy. The coin's rotational kinetic energy equals ½ $I\omega^2$. We measure ω with respect to the coin's axis of rotation, perpendicular to the center of the coin.

Its linear kinetic energy equals ½ mv_{CM}^2. The "CM" subscript indicates that the point used in calculating the linear speed is the coin's center of mass. As Equation 1 shows, the sum of these two types of kinetic energy equals the total kinetic energy.

When an object rolls without slipping, it is often useful know the relationship between its linear and angular velocities. Equation 2 shows this relationship. As the rolling object with radius r rotates through an angle θ, an arc of length $r\theta$ makes contact with the ground. This means the object moves linearly the same distance $r\theta$. Its linear speed v_{CM} is that distance divided by t, or $r\theta/t$. Since θ/t equals ω, we can also say that v_{CM} equals $r\omega$.

The kinetic energy equation and the relationship in Equation 2 are both used to solve the example problem.

equation 1

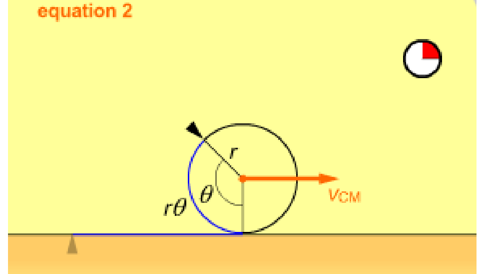


Kinetic energy of rolling object

$$KE = ½mv_{CM}^2 + ½I\omega^2$$

m = mass
v_{CM} = velocity of center of mass
I = moment of inertia
ω = angular velocity

equation 2



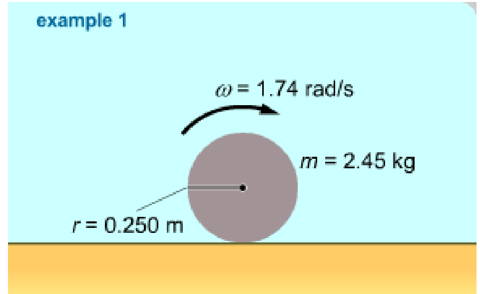
Rolling without slipping: linear and angular velocity

$$v_{CM} = r\theta/t = r\omega$$

v_{CM} = linear velocity of center of mass

r = radius
ω = angular velocity

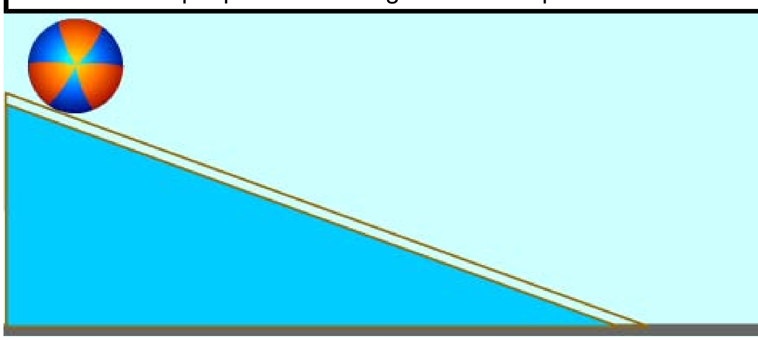
example 1



The solid, uniform disk rolls without slipping. What is its total kinetic energy?

$KE = \frac{1}{2}mv_{CM}^2 + \frac{1}{2}I\omega^2$
$v_{CM} = r\omega = (0.250\text{ m})(1.74\text{ rad/s})$
$v_{CM} = 0.435\text{ m/s}$
$\frac{1}{2}mv_{CM}^2 = 0.232\text{ J}$
$I = \frac{1}{2}mr^2$
$I = \frac{1}{2}(2.45\text{ kg})(0.250\text{ m})^2$
$I = 0.0766\text{ kg·m}^2$
$\frac{1}{2}I\omega^2 = 0.116\text{ J}$
$KE = 0.232\text{ J} + 0.116\text{ J}$
$KE = 0.348\text{ J}$

11.16 - Sample problem: rolling down a ramp

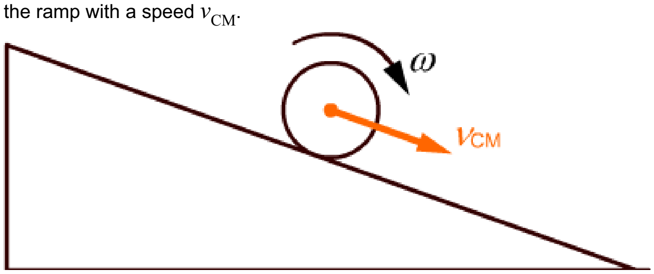


A hollow beach ball of mass 0.10 kg rolls down a ramp without slipping. At the top, its potential energy is 0.75 J. At the bottom, it has no potential energy. What are its linear and rotational kinetic energies there?

As the ball rolls down the ramp, its potential energy is converted to kinetic energy, both linear and rotational.

Draw a diagram

We draw a diagram of the ball partway down the ramp. Since it is rotating, it has an angular velocity ω. Its center of mass is moving parallel to the ramp with a speed v_{CM}.



Variables

The radius of the ball is not stated in the problem. It will cancel out and not be a factor in the answer.

initial (potential) energy	$PE = 0.75$ J
mass	$m = 0.10$ kg
radius	r
angular velocity	ω
speed of center of mass	v_{CM}
moment of inertia	I

What is the strategy?

1. Use the equation for the moment of inertia of a hollow sphere to write an expression for the rotational kinetic energy at any instant in terms of mass, radius and angular velocity.
2. Write an expression for the linear kinetic energy at any instant, also in terms of mass, radius and angular velocity.
3. Find the ratio of the two kinetic energies. This does not change as the ball moves.
4. The total energy at the top of the ramp is all potential energy. At the bottom of the ramp, all the energy is kinetic energy. Distribute the total energy between linear and rotational kinetic energy, according to the constant ratio you just calculated.

Physics principles and equations

We use the conservation of energy. In this case, all the energy at the top of the ramp is potential, and all the energy at the bottom is kinetic.

$$PE \text{ (top)} = KE_L + KE_R \text{ (bottom)}$$

The equations for linear and rotational kinetic energy

$$KE_L = ½\, mv_{CM}^2,\ KE_R = ½\, I\omega^2$$

The relationship between the linear velocity of the ball's center of mass and its angular velocity

$$v_{CM} = r\omega$$

The moment of inertia of a hollow sphere

$$I = \frac{2}{3}mr^2$$

Step-by-step solution

We use the moment of inertia of a hollow sphere to find an expression for the rotational kinetic energy.

Step	Reason
1. $KE_R = ½\, I\omega^2$	rotational kinetic energy
2. $I = 2mr^2/3$	moment of inertia
3. $KE_R = ½(2mr^2/3)\omega^2$	substitute equation 2 into equation 1
4. $KE_R = mr^2\omega^2/3$	simplify

Now we write an expression for the linear kinetic energy.

Step	Reason
5. $KE_L = ½\, mv_{CM}^2$	linear kinetic energy
6. $v_{CM} = r\omega$	rolling without slipping
7. $KE_L = ½\, m(r\omega)^2$	substitute equation 6 into equation 5
8. $KE_L = mr^2\omega^2/2$	simplify

The expressions for kinetic energy in steps 4 and 8 are the same except for a constant factor. We can write the linear kinetic energy as a constant times the rotational kinetic energy.

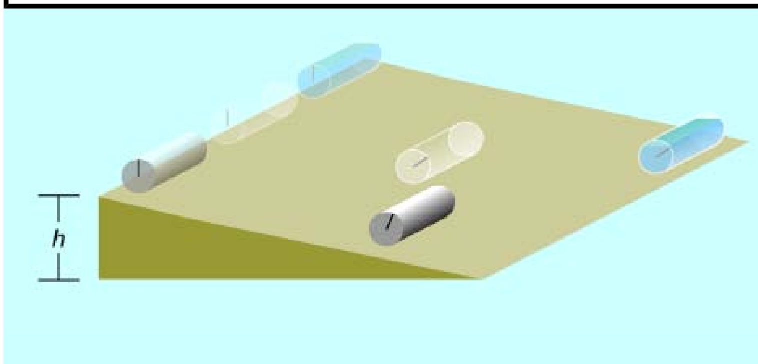
Step	Reason
9. $KE_L/KE_R = (mr^2\omega^2/2)/(\, mr^2\omega^2/3)$	divide equation 8 by equation 4
10. $KE_L = 1.5KE_R$	simplify

Now, we apply the conservation of energy.

Step	Reason
11. $PE = KE_L + KE_R$	conservation of energy
12. $0.75 = 1.5KE_R + KE_R$	given value and substitution from equation 10
13. $KE_R = 0.30$ J	solve
14. $KE_L = 0.45$ J	using equation 10

The ratio of rotational to linear kinetic energy depends on the configuration of the object. For instance, with a hoop, 50% of the kinetic energy is rotational and 50% is linear. For a hollow sphere, as we see in this example, 40% is rotational and 60% is linear. For a given rolling speed, a larger fraction of a hoop's total KE is in the form of rotational KE, because its moment of inertia for a given mass is higher than that of the hollow sphere.

11.17 - Sample problem: rolling cylinders



Three cylinders of the same overall mass roll down an inclined plane: a solid cylinder; a hollow empty cylinder; and a hollow cylinder filled with water. The water has much greater mass than the cylinder enclosing it. All the cylinders have the same radius. Which cylinder has the greatest linear speed at the bottom of the ramp?

The question asks you to determine the linear speeds of three rolling cylinders with the same overall mass but different properties. You may be surprised to learn that the speeds are different!

When doing the calculations for the water filled cylinder, we will assume that the cylinder rotates freely around the water. That is, the water does not rotate. To hold in the water, this cylinder must have end caps, but we will assume their moments of inertia are negligible and ignore this factor. We assume that all the cylinders roll without slipping.

Variables

mass of solid and hollow cylinders	m
mass of water-filled cylinder	m_f
mass of water	m_w
radius of cylinders	r
height of ramp	h
initial energy (at top)	E_i
final energy (at bottom)	E_f
speed of center of mass at bottom	v_{CM}
angular velocity	ω
moment of inertia of a solid cylinder	I_s
moment of inertia of a hollow cylinder	I_h

What is the strategy?

1. The potential energy at the top of the ramp equals the kinetic energy at the bottom. Using conservation of energy, find a general expression for the speed of the center of mass of an object at the bottom of the ramp.
2. For each of the three cylinders, use the moment of inertia formulas to write an expression for the speed of the cylinder.

Physics principles and equations

Gravitational potential energy

$$PE = mgh$$

Kinetic energy of a rolling object

$KE = ½mv^2 + ½I\omega^2$

Moments of inertia

$I = mr^2$ (hollow cylinder)

$I = ½mr^2$ (solid cylinder)

The relationship of the speed and angular velocity for a rolling object

$v_{CM} = r\omega$

We apply the principle of the conservation of energy.

Step-by-step solution

We start by finding a general equation for the speed of an object that has rolled down the ramp, starting from rest at height h. We define the potential energy to be zero at the bottom of the ramp.

Step	Reason
1. $E_i = E_f$	conservation of energy
2. $E_i = PE = mgh$	initial energy is all potential
3. $E_f = KE = ½mv_{CM}^2 + ½I\omega^2$	energy at bottom of ramp is all kinetic
4. $mgh = ½mv_{CM}^2 + ½I\omega^2$	substitute equations 2 and 3 into equation 1
5. $\omega = v_{CM}/r$	relation of angular velocity and speed of center of mass
6. $mgh = ½mv_{CM}^2 + ½I(v_{CM}/r)^2$	substitute equation 5 into equation 4
7. $v_{CM} = \sqrt{\frac{2mgh}{\left(m + \frac{I}{r^2}\right)}}$	solve for v_{CM}

Solid cylinder. Now that we have a general equation for the speed of an object at the bottom of the ramp, we can apply the moment of inertia formulas to find equations for the speeds of the cylinders. We start with the solid cylinder.

Step	Reason
8. $I_s = ½mr^2$	moment of inertia for a solid cylinder
9. $v_{CM} = \sqrt{\frac{2mgh}{\left(m + \frac{\frac{1}{2}mr^2}{r^2}\right)}}$	substitute equation 8 into equation 7
10. $v_{CM} = \sqrt{\frac{4}{3}gh}$	simplify

Hollow cylinder. Notice that the expression for the speed of the solid cylinder is independent of its mass and radius. This means that any solid cylinder will have the same speed as it rolls down a ramp of the same height. Next, we consider the speed of the hollow cylinder.

Step	Reason
11. $I_h = mr^2$	moment of inertia for a hollow cylinder
12. $v_{CM} = \sqrt{\frac{2mgh}{\left(m + \frac{mr^2}{r^2}\right)}}$	substitute equation 11 into equation 7
13. $v_{CM} = \sqrt{gh}$	simplify

Water-filled cylinder. Like the solid cylinder, the speed of the hollow cylinder is independent of its mass and radius. The hollow cylinder rolls slower than the solid cylinder. Finally, we consider the water-filled cylinder.

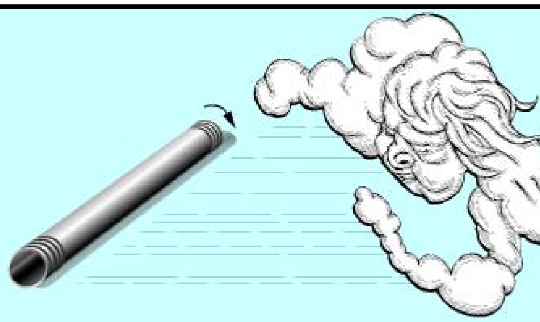
Step	Reason
14. $v_{CM} = \sqrt{\dfrac{2(m_f + m_w)gh}{\left(m_f + m_w + \dfrac{I_h}{r^2}\right)}}$	substitute total mass into equation 7
15. $v_{CM} = \sqrt{\dfrac{2(m_f + m_w)gh}{\left(m_f + m_w + \dfrac{m_f r^2}{r^2}\right)}}$	substitute equation 11 into equation 14
16. $v_{CM} = \sqrt{\dfrac{2(m_f + m_w)gh}{(2m_f + m_w)}}$	simplify
17. $v_{CM} = \sqrt{2gh}$	$m_w >> m_f$

The speed for the water-filled cylinder is the same as an object sliding down the ramp without friction, which makes sense: The cylinder allows the water to slide down the ramp "friction-free." This cylinder has the greatest speed at the bottom of the ramp, because

$$\sqrt{2gh} > \sqrt{\tfrac{4}{3}gh} > \sqrt{gh}$$

If the three cylinders were in a race, the water-filled one would win, the solid one would come in second and the hollow one would be last.

11.18 - Interactive checkpoint: rolling to a stop



A 3.50 kg section of thin-walled pipe with radius 0.110 m is rolling without slipping on a flat horizontal surface. A wind begins to blow with a force of 9.00 N directly against the pipe's direction of movement. If the pipe is rotating at 12.0 radians per second when the wind begins to blow, how far will it roll into the wind before stopping?

Answer:

$\Delta x =$ ______ m

11.19 - Physics at play: a yo-yo

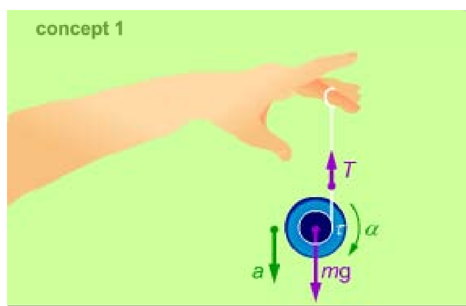
You can use your knowledge of physics to analyze the motion of a yo-yo. Doing so requires several steps, but it is a good way to practice working with forces, torque and energy. To simplify the analysis, we assume that the yo-yo's string has a negligible thickness and mass.

A moving yo-yo functions by converting gravitational potential energy to rotational and linear kinetic energy. To put this more succinctly: As it falls, it rotates. Because of the rotation, the linear acceleration and velocity of the yo-yo are not the same as those of an object in freefall, because its potential energy is converted into both rotational and linear KE.

Now let's analyze yo-yo motion in more detail. Our goal is to create an equation that describes the yo-yo's acceleration as it falls.

Variables

In this analysis we are only concerned with the magnitude of the downward acceleration of the yo-yo, and the magnitude of its angular acceleration.



Yo-yo physics
Net force = weight – tension
String tension creates torque that causes rotation

For this reason we abandon our usual convention and take downward as the positive linear direction.

mass of yo-yo	m
acceleration of yo-yo	a
acceleration of gravity	$g = 9.80\ \text{m/s}^2$
upward force exerted by string	T
torque of string on yo-yo spindle	τ
radius of spindle	r
moment of inertia of yo-yo	I
angular acceleration of yo-yo	α

Strategy

1. Analyze the linear motion of the yo-yo using Newton's second law. The equation will contain the tension T in the yo-yo string.
2. Analyze the angular motion of the yo-yo using the rotational form of Newton's second law. This equation, too, will contain T.
3. Combine the equations for linear and rotational motion, using the common variable T, and solve for the downward acceleration of the descending yo-yo.

Physics principles and equations

We will use two versions of Newton's second law to calculate linear and angular acceleration.

$$\Sigma F = ma, \quad \Sigma\tau = I\alpha$$

For a force perpendicular to the line from the axis of rotation to the point where the force is applied, the torque is

$$\tau = rF$$

The resulting tangential acceleration is

$$a = r\alpha$$

Step-by-step derivation

In the first steps we analyze the linear motion of the yo-yo. Note that the equation in the third step contains the tension force exerted by the string.

Step	Reason
1. $\Sigma F = ma$	Newton's second law
2. $\Sigma F = mg + (-T)$	inspection
3. $mg - T = ma$	substitute equation 2 into equation 1

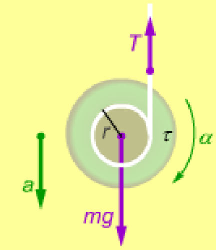
We now analyze the angular acceleration of the yo-yo using Newton's second law for rotation. We solve the resulting equation for the tension force exerted by the string.

Step	Reason
4. $\Sigma\tau = I\alpha$	Newton's second law for rotation
5. $\Sigma\tau = rT$	torque and tangential force
6. $\alpha = a/r$	relationship of angular and tangential acceleration
7. $rT = I(a/r)$	substitute equations 5 and 6 into equation 4
8. $T = Ia/r^2$	solve for T

Now we combine the linear and rotational analyses and solve for the linear acceleration.

Step	Reason
9. $mg - Ia/r^2 = ma$	substitute equation 8 into equation 3
10. $a = \dfrac{g}{1+(I/mr^2)}$	solve for a

equation 1



Acceleration of descending yo-yo

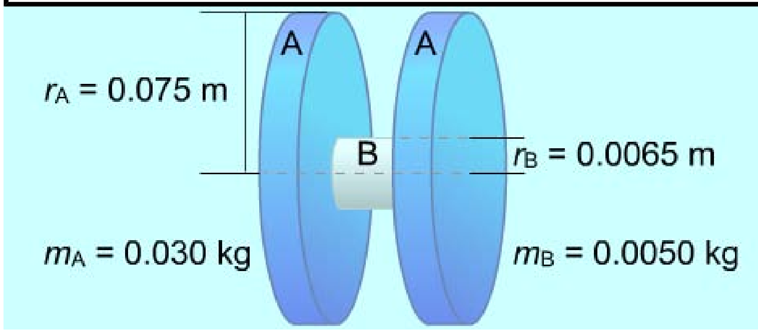
$$a = \frac{g}{1+(I/mr^2)}$$

a = acceleration of yo-yo
g = acceleration of gravity
I = moment of inertia of yo-yo
m = mass of yo-yo
r = radius of spindle

We now have an equation for the linear acceleration of a yo-yo based on its moment of inertia, its mass, and the radius of its spindle. All three depend on the yo-yo's geometry and composition.

"Sleeper" yo-yos, and especially those intended for competition, work best if they acquire a great deal of rotational KE as they fall. The faster they rotate, the more gyroscopically stable they are for tricks, and the longer they rotate. The rotational KE is greater when less of the original potential energy is converted to linear KE, so a slower-descending yo-yo is better. If you inspect the formula, you will see that a can be reduced for a yo-yo with a given moment of inertia and mass by making the spindle radius r as small as possible. If you inspect competition yo-yos, you will see that their spindles are quite slender.

11.20 - Sample problem: acceleration of a yo-yo



The yo-yo consists of two identical solid cylindrical disks and an inner spindle. What is the magnitude of acceleration of the yo-yo as it falls?

The section on the physics of yo-yos states an equation for the acceleration based on the yo-yo's moment of inertia, mass and the radius of its spindle. The spindle radius is given, and the mass of the yo-yo can be easily calculated. The main task in solving the problem is determining the yo-yo's moment of inertia.

Variables

mass of disk A	$m_A = 0.030$ kg
mass of spindle	$m_B = 0.0050$ kg
total mass	$m = 2m_A + m_B$
radius of disk A	$r_A = 0.075$ m
radius of spindle	$r_B = 0.0065$ m
moment of inertia, disk A	I_A
moment of inertia, spindle	I_B
total moment of inertia	I

What is the strategy?

1. Calculate the moment of inertia of each of the cylinders that make up the yo-yo, the spindle and the two identical disks.
2. Calculate the total moment of inertia of the yo-yo.
3. Use the equation for the acceleration of a yo-yo to solve the problem.

Physics principles and equations

The three rotating cylinders that make up the yo-yo share a common axis of rotation. This means we can add their moments to calculate the total moment of inertia.

The moment of inertia of a solid cylinder rotating around its central axis

$$I = \tfrac{1}{2} mr^2$$

The equation for the acceleration of a yo-yo is

$$a = \frac{g}{1 + (I/mr^2)}$$

where g is gravitational acceleration.

Steps

First, calculate the moment of inertia of each of the cylinders making up the yo-yo. We start with the identical disks A, and then calculate the moment of the spindle B.

Step	Reason
1. $I_A = \frac{1}{2} m_A r_A^2$	moment of inertia of solid cylinder
2. $I_A = \frac{1}{2}(0.030\ \text{kg})(0.075\ \text{m})^2$ $I_A = 8.44 \times 10^{-5}\ \text{kg} \cdot \text{m}^2$	enter values and multiply
3. $I_B = \frac{1}{2} m_B r_B^2$ $I_B = \frac{1}{2}(0.0050\ \text{kg})(0.0065\ \text{m})^2$ $I_B = 1.06 \times 10^{-7}\ \text{kg} \cdot \text{m}^2$	moment of inertia of spindle

Now we calculate the moment of inertia of the yo-yo.

Step	Reason
4. $I = 2I_A + I_B$	yo-yo has two disks and spindle
5. $I = 2(8.44 \times 10^{-5}\ \text{kg} \cdot \text{m}^2) + 1.06 \times 10^{-7}\ \text{kg} \cdot \text{m}^2$ $I = 1.69 \times 10^{-4}\ \text{kg} \cdot \text{m}^2$	sum of moments

Now we use the equation stated above for the acceleration of a yo-yo.

Step	Reason
6. $a = \dfrac{g}{1 + (I / m r_B^2)}$	acceleration of a yo-yo
7. $m = 2m_A + m_B$ $m = 2(0.030\ \text{kg}) + 0.0050\ \text{kg}$ $m = 0.065\ \text{kg}$	total mass
8. $a = \dfrac{9.80\ \text{m/s}^2}{1 + \dfrac{1.69 \times 10^{-4}\ \text{kg} \cdot \text{m}^2}{(0.065\ \text{kg})(0.0065\ \text{m})^2}}$	enter values into equation 6
9. $a = 0.16\ \text{m/s}^2$	evaluate

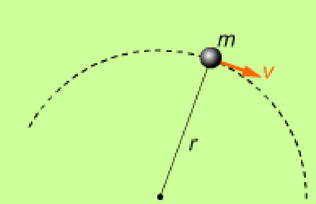
11.21 - Angular momentum of a particle in circular motion

The concepts of linear momentum and conservation of linear momentum prove very useful in understanding phenomena such as collisions. *Angular momentum* is the rotational analog of linear momentum, and it too proves quite useful in certain settings. For instance, we can use the concept of angular momentum to analyze an ice skater's graceful spins.

In this section, we focus on the angular momentum of a single particle revolving in a circle. Angular momentum is always calculated using a point called the origin. With circular motion, the simple and intuitive choice for the origin is the center of the circle, and that is the point we will use here. The letter **L** represents angular momentum.

As with linear momentum, angular momentum is proportional to mass and velocity. However, with rotational motion, the distance of the particle from the origin must be taken into account, as well. With circular motion, the amount of angular momentum equals the product of mass, speed and the radius of the circle: mvr. Another way to state the same thing is to say that the amount of angular momentum equals the linear momentum mv times the radius r.

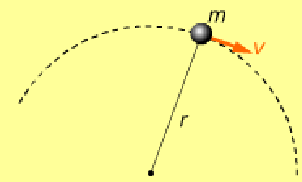
concept 1



Angular momentum of a particle
Proportional to mass, speed, and distance from origin

Like linear momentum, angular momentum is a vector. When the motion is counterclockwise, by convention, the vector is positive. The angular momentum of clockwise motion is negative. The units for angular momentum are kilogram-meter^2 per second (kg·m^2/s).

equation 1

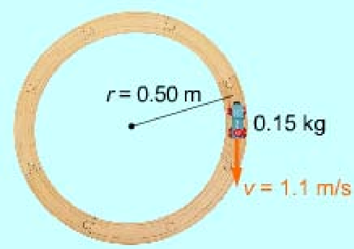


Angular momentum of a particle

$L = mvr$

L = angular momentum
m= mass
v = speed
r = distance from origin (radius)
Counterclockwise +, clockwise –
Units: kg·m²/s

example 1



How much angular momentum does the engine have?

$L = -mvr$
$L = -(0.15\text{ kg})(1.1\text{ m/s})(0.50\text{ m})$
$L = -0.083\text{ kg·m}^2\text{/s}$

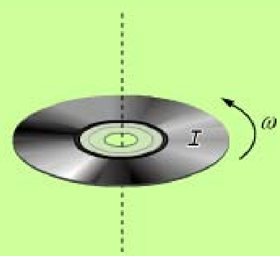
11.22 - Angular momentum of a rigid body

On the right, you see a familiar sight: a rotating compact disc. In the prior section, we defined the angular momentum of a single particle as the product of its mass, speed and radial distance from the axis of revolution. A CD is more complex than that. It consists of many particles rotating at different distances from a common axis of rotation. The CD is rigid, which means the particles all rotate with the same angular velocity, and each remains at a constant radial distance from the axis.

We can determine the angular momentum of the CD by summing the angular momenta of all the particles that make it up. The resulting sum can be expressed concisely using the concept of moment of inertia. The magnitude of the angular momentum of the CD equals the product of its moment of inertia, I, and its angular velocity, ω.

We derive this formula for calculating angular momentum below. In Equation 1, you see one of the rotating particles drawn, with its mass, velocity and radius indicated.

concept 1



Angular momentum of a rigid body
Product of moment of inertia, angular velocity

Variables

mass of a particle	m_i
tangential (linear) speed of a particle	v_i
radius of a particle	r_i
angular momentum of particle	L_i
angular momentum of CD	L
angular velocity of CD	ω
moment of inertia of CD	I

Strategy

1. Express the angular momentum of the CD as the sum of the angular momenta of all the particles of mass that compose it.
2. Replace the speed of each particle with the angular velocity of the CD times the radial distance of the particle from the axis of rotation.
3. Express the sum in concise form using the moment of inertia of the CD.

Physics principles and equations

The angular momentum of a particle in circular motion

$$L = mvr$$

We will use the equation that relates tangential speed and angular velocity.

$$v = r\omega$$

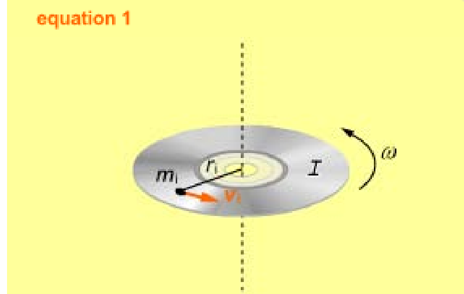
The formula for the moment of inertia of a rotating body

$$I = \Sigma m_i r_i^2$$

Step-by-step derivation

First, we express the angular momentum of the CD as the sum of the angular momenta of the particles that make it up.

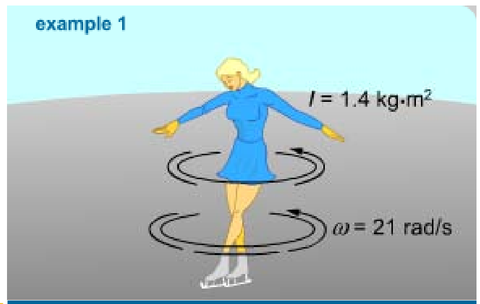
Step	Reason
1. $L_i = m_i v_i r_i$	definition of angular momentum
2. $L = \Sigma m_i v_i r_i$	angular momentum of object is sum of particles



Angular momentum of a rigid body

$$L = I\omega$$

L = angular momentum
I = moment of inertia
ω = angular velocity



How much angular momentum does the skater have?

$L = I\omega$
$L = (1.4\ \text{kg·m}^2)(21\ \text{rad/s})$
$L = 29\ \text{kg·m}^2\text{/s}$

We now express the speed of the ith particle as its radius times the constant angular velocity ω, which we then factor out of the sum. The angular velocity is the same for all particles in a rigid body.

Step	Reason
3. $v_i = r_i\omega$	tangential speed and angular velocity
4. $L = \Sigma m_i(r_i\omega)r_i$	substitute equation 3 into equation 2
5. $L = (\Sigma m_i r_i^2)\omega$	factor out ω

In the final steps we express the above result concisely, replacing the sum in the last equation by the single quantity I.

Step	Reason
6. $I = \Sigma m_i r_i^2$	moment of inertia
7. $L = I\omega$	substitute equation 6 into equation 5

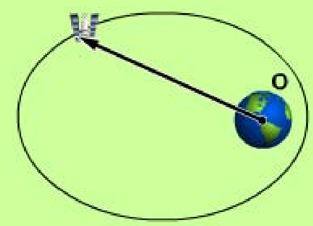
11.23 - Angular momentum: general motion

Angular momentum can be calculated for any object moving with respect to any point designated as the origin. The motion need not be circular and the origin can be at any location. In this section, we apply the concept to a satellite in an elliptical orbit around the Earth, using the Earth as the origin. (The diagram is not drawn to scale.)

To calculate the angular momentum, we use the position vector **r** that points from the origin to the satellite. The angular momentum of a particle with respect to the origin is the cross product of the particle's position vector and its linear momentum. Since the linear momentum is the particle's mass times its velocity, we can also say the angular momentum equals the mass times the cross product of the radius and velocity. These equations are shown on the right. Since angular momentum is a cross product vector, its direction is calculated with the right-hand rule for cross products.

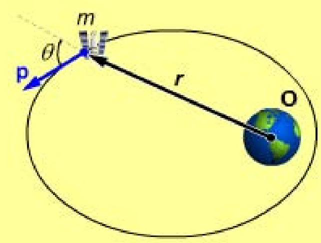
The equations can be used for any origin and motion along any path: an ellipse, a straight line, and so on. Circular motion, with the origin at the center, is a special case for this equation: Because the particle's velocity is always tangent to the radius, the angle is $90°$ and the sine is one. This means, as was stated earlier, that the particle's angular momentum in this case is equal to the product of its mass, speed and the radius of its circular path.

concept 1



Angular momentum: general case
Measured with respect to an origin

equation 1

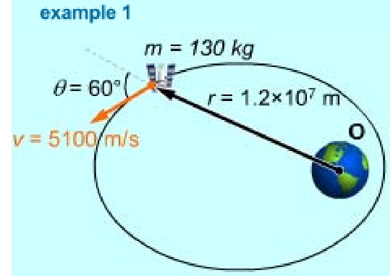


Angular momentum: general case

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = m(\mathbf{r} \times \mathbf{v})$$
$$L = mvr \sin \theta$$

$\mathbf{L}$ = angular momentum
$\mathbf{r}$ = position vector
$\mathbf{p}$ = linear momentum
m = mass
$\mathbf{v}$ = velocity
θ = smaller angle between $\mathbf{r}$ and $\mathbf{p}$

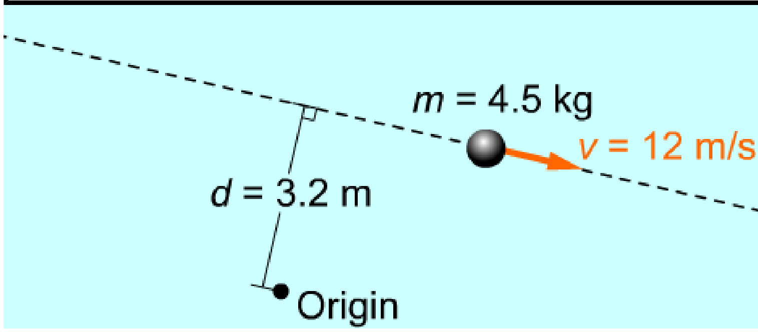
example 1



What is the angular momentum of the satellite with the Earth as the origin?

$L = mvr \sin \theta$
$L = (130)(5100)(1.2\times10^7)(\sin 60°)$
$\mathbf{L} = 6.9\times10^{12}$ kg·m²/s, out of screen

11.24 - Sample problem: object moving in a straight line

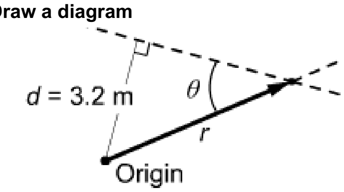


What is the magnitude of angular momentum of the ball?

You are asked to calculate the angular momentum of a ball moving at a constant velocity, subject to no external forces, in a straight line. The illustration shows the ball's path and the shortest distance d between the path and the origin. Calculating the angular momentum of an object moving in a straight line may seem unusual, but it is useful in some situations.

One facet of the solution is quite significant: The angular momentum of the ball is constant. This is shown in Step 5 below. It is a function of the location of the origin, but for any origin you choose, the ball will have a constant angular momentum, equal to its linear momentum times d. If you anticipated this, you correctly anticipated that when there is no net external torque, angular momentum remains constant. (This resembles a precept in linear momentum: When there is no net external force, linear momentum remains constant.)

Draw a diagram



Variables

distance from origin to path of motion	$d = 3.2$ m
velocity	$v = 12$ m/s
mass	$m = 4.5$ kg
distance from origin	r
angle between position vector and path	θ
momentum	L

What is the strategy?

1. The equation that defines angular momentum uses the length r of the position vector from the origin to the object and the angle θ between the position vector and the object's path. Find an expression for r in terms of the angle θ and the distance d.
2. Use this expression in the general equation for angular momentum, and simplify. This yields an equation that can be used to solve the problem.
3. Enter the known values and do the computation.

Physics principles and equations

The general equation for the magnitude of angular momentum of a particle is

$$L = mvr \sin \theta$$

Step-by-step solution

We start by finding an expression for r in terms of d and θ. This will prove useful in the next set of steps.

Step	Reason
1. $\sin \theta = d/r$	from diagram
2. $r = \dfrac{d}{\sin\theta}$	solve for r

Now we use the equation for angular momentum, and the expression for r we just derived. This will give us an equation for angular momentum without r or θ.

Step	Reason
3. $L = mvr \sin\theta$	definition of angular momentum
4. $L = mv\left(\dfrac{d}{\sin\theta}\right)\sin\theta$	substitute step 2 into step 3
5. $L = mvd$	cancel sin θ factors
6. $L = (4.5\text{ kg})(12\text{ m/s})(3.2\text{ m})$ $L = 170\text{ kg}\cdot\text{m}^2/\text{s}$	evaluate

11.25 - Comparison of rotational and linear motion

In the table, we summarize the correspondence between linear and rotational motion. Seeing how these concepts relate may help you remember them.

linear		rotational	
x	position	θ	angular position
Δx	displacement	$\Delta\theta$	angular displacement
$v = \Delta x/\Delta t$	velocity	$\omega = \Delta\theta/\Delta t$	angular velocity
$a = \Delta v/\Delta t$	acceleration	$\alpha = \Delta\omega/\Delta t$	angular acceleration
F	force	τ	torque
m	mass	I	moment of inertia
$\Sigma F = ma$	Newton's 2nd law	$\Sigma\tau = I\alpha$	Newton's 2nd law (rotation)
$\frac{1}{2}mv^2$	kinetic energy	$\frac{1}{2}I\omega^2$	kinetic energy
$F\Delta x$	work (force parallel)	$\tau\Delta\theta$	work
$p = mv$	momentum	$L = I\omega$	angular momentum

11.26 - Torque and angular momentum

A net force changes an object's velocity, which means its linear momentum changes as well. Similarly, a net torque changes a rotating object's angular velocity, and this changes its angular momentum.

As Equation 1 shows, the product of torque and an interval of time equals the change in angular momentum. This equation is analogous to the equation from linear dynamics stating that the impulse (the product of average force and elapsed time) equals the change in linear momentum.

In the illustration to the right, we show a satellite rotating at a constant angular velocity and then firing its thruster rockets, which changes its angular momentum. The thrusters apply a constant torque $\boldsymbol{\tau}$ to the satellite for an elapsed time Δt. In Example 1, the satellite's change in angular momentum is calculated using the equation. (The change in the satellite's mass and moment of inertia resulting from the expelled fuel are small enough to be ignored.)

equation 1

Torque and angular momentum

$$\boldsymbol{\tau}\Delta t = \Delta\mathbf{L}$$

$\boldsymbol{\tau}$ = torque
Δt = time interval
$\Delta\mathbf{L}$ = change in angular momentum

example 1

The rockets provide 56 N·m of torque for 3.0 s. What is the amount of change in the satellite's angular momentum?

$\tau\Delta t = \Delta L$

$(56 \text{ N·m})(3.0 \text{ s}) = \Delta L$
$\Delta L = 168 \text{ kg·m}^2/\text{s}$

11.27 - Conservation of angular momentum

concept 1

Angular momentum conserved
No external torque
Angular momentum is constant

Linear momentum is conserved when there is no external net force acting on a system. Similarly, angular momentum is conserved when there is no net external torque. To put it another way, if there is no net external torque, the initial angular momentum equals the final angular momentum. This is stated in Equation 1.

The principle of conservation of linear momentum is often applied to collisions, and the masses of the colliding objects are assumed to remain constant. However, with angular momentum, we often examine what occurs when the moment of inertia of a body changes. Since angular momentum equals the product of the moment of inertia and angular velocity, if one of these properties changes, the other must as well for the angular momentum to stay the same. This principle is used both in classroom demonstrations and in the world of sports. In a common classroom demonstration, a student is set rotating on a stool. The student holds weights in each hand, and as she pushes them away from her body, she slows down. In doing so, she demonstrates the conservation of angular momentum: As her moment of inertia increases, her angular velocity decreases. In contrast, pulling the weights close in to her body decreases her moment of inertia and increases her angular velocity.

Ice skaters apply this principle skillfully. When they wish to spin rapidly, they wrap their arms tightly around their bodies. They decrease their moment of inertia to increase their angular velocity. You can see images of a skater applying this principle to the right and above.

equation 1

Conservation of angular momentum

$L_i = L_f$
$I_i\omega_i = I_f\omega_f$

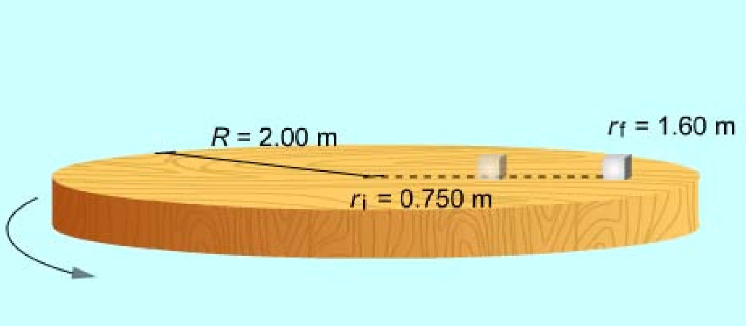
L = angular momentum
I = moment of inertia
ω = angular velocity

example 1

The skater pulls in her arms, cutting her moment of inertia in half. How much does her angular velocity change?

$I_i\omega_i = I_f\omega_f$
$I_i\omega_i = (\frac{1}{2}I_i)\omega_f$
$\omega_f = 2\omega_i$ (it doubles)

11.28 - Interactive checkpoint: a rotating disk

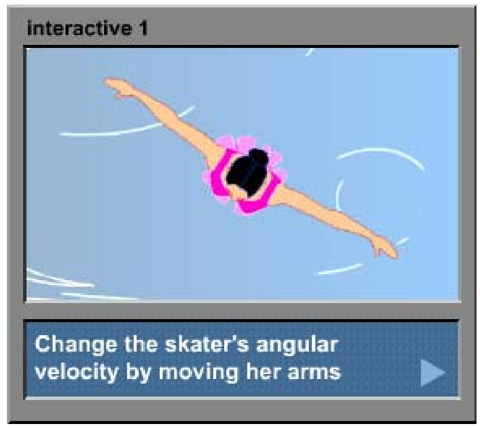


A 45.0 kg cube is resting 0.750 m from the center of a solid 125 kg disk that is rotating at 2.80 rad/s. The radius of the disk is 2.00 m. The cube slides radially to a location 1.60 m from the center of the disk, changing the moment of inertia of the system. What is the new angular velocity of the system? Treat the cube as a point mass in your calculation.

Answer:

ω_f = ______ rad/s

11.29 - Interactive summary problem: dynamics of skating



In the simulations on the right, you are a champion ice skater in the final round of the toughest competition of your career. You want to get the high score and win the gold medal. In addition to your grace, athleticism, charm, coach, supportive parents, agent, diet, endorsements and hard work, you will use your knowledge of physics to take home the top prize.

In the first simulation, you are in the final spin of your routine, with your arms fully extended. You want to increase your angular velocity solely by changing the position of your arms. Each arm has a mass of 1.80 kilograms. For simplicity's sake, consider each arm as a point of mass at an effective distance of 0.480 meters away from the vertical center of your body, your axis of rotation. The arms move symmetrically in the simulation.

(We stress "point of mass" because your arms contain mass at different distances from the axis of rotation, and since they are not uniform, calculating their moments of inertia would be complex. If you like, you could imagine that you are holding a weight in each hand, and the only thing you change is the position of those weights. Of course, we would then ask you to treat those weights as point masses…)

Your initial angular velocity is 2.00π rad/s. You want to increase your angular velocity to 4.50π rad/s. Your initial moment of inertia is 1.40 kg·m^2. The moment of inertia of the rest of your body remains the same during this maneuver.

Enter a distance for your final arm position to the nearest 0.001 m to change your moment of inertia and angular velocity, and press GO. Press RESET to start over.

If you have difficulty solving this problem, review the sections on moment of inertia, angular momentum, and conservation of angular momentum. Remember that the moments of inertia of a complex object can be added.

In the second simulation, you are initially spinning at 4.00π rad/s, and it is time to stop your movement and face the judges. To stop, you place the tip of one skate on the ice, creating a frictional force. You apply the skate tip to the ice when you are facing away from the judges, and you want to stop in exactly 2.50 turns so you finish up facing them. The skate tip is 0.200 meters from the axis of rotation, and the frictional force is always perpendicular to the radius from the axis of rotation. Your moment of inertia during this maneuver is 0.601 kg·m^2.

Determine the magnitude of the frictional force needed to stop in 2.50 turns. To do this, determine the angular acceleration to reach zero angular velocity in 2.50 turns and then use the relationship of torque, moment of inertia and angular acceleration. Enter the force you calculate to the nearest 0.1 N and press GO to start the simulation. You will see the angular acceleration in an output gauge. Press RESET to start over.

To solve this problem, you may want to review the section on torque and angular acceleration. You will also need to use one of the rotational motion equations.

11.30 - Gotchas

A torque is a force. No, it is not. A net torque causes **angular** acceleration. It requires a force.

A torque that causes counterclockwise acceleration is a positive torque. Yes, and a torque that causes clockwise acceleration is negative.

I have a baseball bat. I shave off some weight from the handle and put it on the head of the bat. A baseball player thinks I have changed the bat's moment of inertia. Is he right? Yes. A baseball player swings from the handle, so you have increased the amount of mass at the farthest distance, changing the bat's moment. The player will find it harder to apply angular acceleration to the bat.

A skater begins to rotate more slowly, so his angular momentum must be changing. Not necessarily. An external torque is required to change the angular momentum. The slower rotation could instead be caused by the skater altering his moment of inertia, perhaps by moving his hands farther from his body. On the other hand, if he digs the tip of a skate into the ice, that torque would reduce his angular momentum.

11.31 - Summary

Torque is a force that causes rotation. Torque is a vector quantity with units N·m.

A useful quantity in determining torque is the lever arm. It is the perpendicular distance from the line of action of the force to the axis of rotation. Torque can be increased by having a longer lever arm or a stronger force.

Torque is calculated as the cross product of two vectors. The magnitude of the cross product is the product of the multiplied vectors' magnitudes and the sine of the smaller angle between them. To determine the direction of the cross product, use the right-hand rule.

An object's moment of inertia is a measure of its resistance to angular acceleration, just as an object's mass is a measure of its resistance to linear acceleration. The moment of inertia is measured in kg·m^2 and depends not only upon an object's mass, but also on the distribution of that mass around the axis of rotation. The farther the distribution of the mass from the axis, the greater the moment of inertia.

Another linear analogy applies: Just as Newton's second law states that net force equals mass times linear acceleration, the net torque on an object equals its moment of inertia times its angular acceleration.

The parallel axis theorem lets you calculate the moment of inertia of an object rotating about an axis that is parallel to an axis through its center of mass. This equation is helpful when the moment of inertia about the center of mass is already known.

Work done by a torque equals the product of the torque and the angular displacement when the torque is constant.

Rotational kinetic energy depends upon the moment of inertia and the angular velocity. Mechanical devices called flywheels can store rotational kinetic energy.

To find the total kinetic energy of an object that rolls as it rotates, you add its linear kinetic energy to its rotational kinetic energy.

Angular momentum is the rotational analog to linear momentum. Its units are kg·m^2/s. The magnitude of the angular momentum of an object in circular motion is the product of its mass, tangential velocity, and the radius of its path. The angular momentum of a rigid rotating body equals its moment of inertia multiplied by its angular velocity.

In general, angular momentum is the cross product of position and linear momentum.

Just as a change in linear momentum (impulse) is equal to a force times its duration, a change in angular momentum is equal to a torque times its duration.

Angular momentum is conserved in the absence of a net torque on the system.

Equations

Torque

$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$

$\tau = rF \sin \theta$

Newton's second law for rotation

$\Sigma\tau = I\alpha$

Moment of inertia

$I = \Sigma mr^2$

Parallel axis theorem

$I_{\mathrm{P}} = I_{\mathrm{CM}} + Mh^2$

Rotational work

$W = \tau\Delta\theta$

Rotational kinetic energy

$KE = \frac{1}{2} I\omega^2$

Kinetic energy of rolling object

$KE = \frac{1}{2} mv_{\mathrm{CM}}^2 + \frac{1}{2} I\omega^2$

$v_{\mathrm{CM}} = r\theta/t = r\omega$

Angular momentum

$L = I\omega$

$\mathbf{L} = \mathbf{r} \times \mathbf{p} = m(\mathbf{r} \times \mathbf{v})$

$L = mvr \sin \theta$

$\tau\Delta t = \Delta L$

Conservation of angular momentum

$L_{\mathrm{i}} = L_{\mathrm{f}}$

$I_{\mathrm{i}}\omega_{\mathrm{i}} = I_{\mathrm{f}}\omega_{\mathrm{f}}$

Conceptual Problems

C.1 You get a flat tire. You are trying to remove one of the lug nuts from the hubcap with a wrench. No matter how hard you pull, the nut will not budge. You examine the contents of your car to see if you have anything that will be useful. Which of the following objects can best help you remove the nut: a long, strong piece of rope; a long, hollow pipe; road flares; a first aid kit; 2 wool blankets; an ice scraper; your pet toy poodle. Explain.

C.2 (a) Can an object have an angular velocity if there is no net torque acting on it? (b) Can an object have a net torque acting on it if it has zero angular velocity?

(a) ○ Yes ○ No
(b) ○ Yes ○ No

C.3 On which of the following does the moment of inertia depend?

- ☐ Angular velocity
- ☐ Angular momentum
- ☐ Shape of the object
- ☐ Location of axis of rotation
- ☐ Mass
- ☐ Linear velocity

C.4 An object of fixed mass and rigid shape has one unique value for its moment of inertia. True or false? Explain your answer.

○ True ○ False

C.5 Compared to a solid sphere, will a hollow spherical shell (like a basketball) of the same mass and radius have a greater or a lesser moment of inertia for rotations about an axis passing through the center? Explain your answer.

○ Greater ○ Lesser

C.6 Scoring in the sport of gymnastics has much to do with the difficulty of the skills performed, with higher scores awarded to more difficult tricks. The position of a gymnast's body in a trick determines the difficulty. There are three positions to hold the body in a flip: the tuck, in which the legs are bent and tucked into the body; the pike, in which the body is bent at the hips but the legs remain straight; and the layout, in which the whole body is straight. A layout receives a higher score than a pike which receives a higher score than a tuck. Use the principles of rotational dynamics to describe why the position of the body affects the difficulty of a flip.

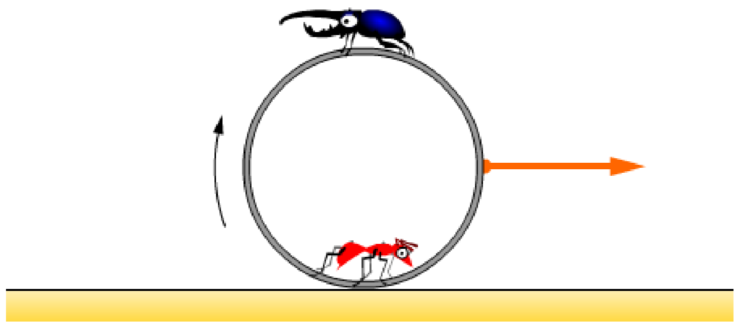
C.7 A bicyclist rides down the street. From her point of view, what is the direction of the angular momentum vector due to her wheel rotation?

i. Straight up
ii. Straight down
iii. Forward
iv. Backward
v. To her left
vi. To her right

C.8 In extreme motocross events, daring though perhaps slightly lunatic motorcycle riders launch themselves from dirt ramps, then perform various acrobatic stunts before landing the vehicle on a second ramp. After the initial launch, the bike is aimed skyward, but for safety reasons, the wheels should both touch down at the same time upon the other downward-sloping ramp, which means that the body of the motorcycle must rotate while the bike is in flight. The rider can rotate the bike about a horizontal axis while it is flying through the air by using the accelerator or the brake, that is, by either speeding up or slowing down the rotation of the rear wheel.
(a) What principle(s) allow the rider to rotate the body of the motorbike while flying through the air? Check all that apply. (b) In order to tip the nose of the motorcycle downward, should the rider use the accelerator or the brake? (c) Explain your answer to part b.

(a) ☐ Conservation of angular momentum
☐ Conservation of linear momentum
☐ Parallel axis theorem
(b) ○ Accelerator ○ Brake
(c)

C.9 A thin, small hoop is rolling slowly along the ground. An ant is walking on the inside of the hoop and a beetle is walking on top of the rolling hoop, in the directions shown. They walk at such a speed that they maintain their positions at the bottom and top of the hoop, respectively. (a) Which insect has to walk at a higher speed? (b) According to a nearby spider sitting on the ground watching the spectacle, which insect is moving faster with respect to the Earth?



(a) i. The beetle on top walks faster
ii. The ant inside walks faster
iii. They walk at the same speed

(b) i. The beetle on top travels faster
ii. The ant inside travels faster
iii. They move at the same speed

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Does changing the distance from the axis of rotation affect the angular acceleration? (b) Where should you push to create the maximum torque and angular acceleration?

(a) ◯ Yes ◯ No

(b) i. Farthest from the axis of rotation
ii. In the middle of the bridge
iii. Closest to the axis of rotation

Section 1 - Torque

1.1 The wheel on a car is held in place by four nuts. Each nut should be tightened to 94.0 N·m of torque to be secure. If you have a wrench with a handle that is 0.250 m long, what minimum force do you need to exert perpendicular to the end of the wrench to tighten a nut correctly?

__________ N

1.2 A 1.1 kg birdfeeder hangs from a horizontal tree branch. The birdfeeder is attached to the branch at a point that is 1.1 m from the trunk. What is the amount of torque exerted by the birdfeeder on the branch? The origin is at the pivot point, where the branch attaches to the trunk.

__________ N · m

1.3 Bob and Ray push on a door from opposite sides. They both push perpendicular to the door. Bob pushes 0.63 m from the door hinge with a force of 89 N. Ray pushes 0.57 m from the door hinge with a force of 98 N, in a manner that tends to turn the door in a clockwise direction. What is the net torque on the door?

__________ N · m

Section 2 - Torque, angle and lever arm

2.1 A 3.30 kg birdfeeder hangs from the tip of a 1.10 m pole that sticks up from the ground at a 65.0° angle. What is the magnitude of the torque exerted on the pole by the birdfeeder? Treat the bottom end of the pole as the pivot point.

__________ N · m

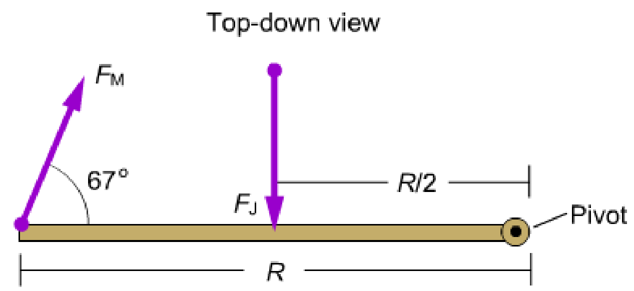
2.2 You want to exert a torque of at least 35.0 N·m on a wrench whose handle is 0.150 m long. If you can provide a force of 355 N to the end of the wrench, what is the minimum angle at which you can apply the force in order to achieve the desired torque?

__________ °

2.3 A simple pendulum consists of a 0.500 kg "bob" (a small but massive object) at the end of a light rod, which is allowed to swing back and forth from a pivot point. The mass of the rod is negligible compared to the mass of the bob. The bob on the pendulum is positioned so the rod makes a 12.0° angle with the vertical. The rod is 2.70 m long. What is the magnitude of the torque exerted on the rod by the weight of the bob?

__________ N · m

2.4 A straight 2.60 m rod pivots around a vertical axis so that it can swing freely in a horizontal plane. Jane pushes perpendicular to the rod at its midpoint with 98.0 N of force, directed horizontally, to create a torque. Matt has attached a rope to the end of the rod, and is pulling on it horizontally to create an opposing torque. The rope creates a 67.0° angle with the rod. With what force should Matt pull so that the net torque on the rod is zero?



__________ N

Section 3 - Cross product of vectors

3.1 Two vectors **A** and **B** form a 57° angle when placed tail-to-tail. **A** has a magnitude of 3.0 and **B** has a magnitude of 5.6. What is the magnitude of their cross product?

____

3.2 At a point on the Earth, vector **A** points due north and vector **B** points due east. Both are horizontal with respect to the ground. (a) What is the direction of **A**×**B**? (b) **B**×**A**?

(a)
- i. North
- ii. West
- iii. South
- iv. East
- v. Up toward the sky
- vi. Down toward the ground

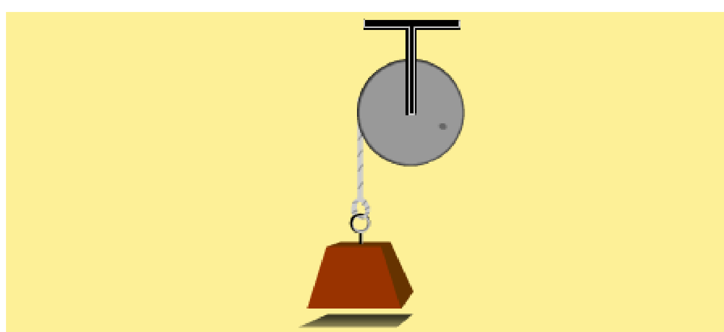
(b)
- i. North
- ii. West
- iii. South
- iv. East
- v. Up toward the sky
- vi. Down toward the ground

3.3 What is the cross product **A** × **B** in rectangular notation if **A** = (2, 2, 0) and **B** = (3, 0, 0)?

(__________ , __________ , __________)

Section 4 - Torque, moment of inertia and angular acceleration

4.1 The pulley shown in the illustration has a radius of 2.70 m and a moment of inertia of 39.0 kg·m^2. The hanging mass is 4.20 kg and it exerts a force tangent to the edge of the pulley. What is the angular acceleration of the pulley?

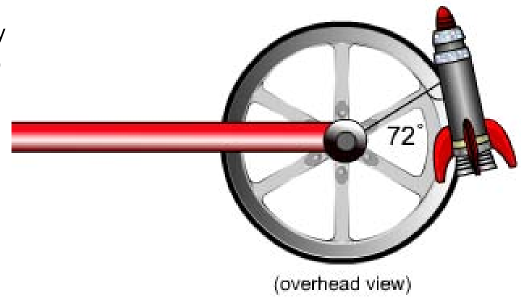


__________ rad/s^2

4.2 A string is wound tight around the spindle of a top, and then pulled to spin the top. While it is pulled, the string exerts a constant torque of 0.150 N·m on the top. In the first 0.220 s of its motion, the top reaches an angular velocity of 12.0 rad/s. What is the moment of inertia of the top?

__________ kg · m^2

4.3 A wheel with a radius of 0.71 m is mounted horizontally on a frictionless vertical axis. A toy rocket is attached to the edge of the wheel, so that it makes a 72° angle with a radius line, as shown. The moment of inertia of the wheel and rocket is 0.14 kg·m^2. If the rocket starts from rest and fires for 8.4 seconds with a constant force of 1.3 N, what is the angular velocity of the wheel when the rocket stops firing?



__________ rad/s

Section 5 - Calculating the moment of inertia

5.1 Four small balls are arranged at the corners of a rigid metal square with sides of length 3.0 m. An axis of rotation in the plane of the square passes through the center of the square, and is parallel to two sides of the square. On one side of the axis, the two balls have masses 1.8 kg and 2.3 kg; on the other side, 1.5 kg and 2.7 kg. The mass of the square is negligible compared to the mass of the balls. What is the moment of inertia of the system for this axis?

__________ kg · m^2

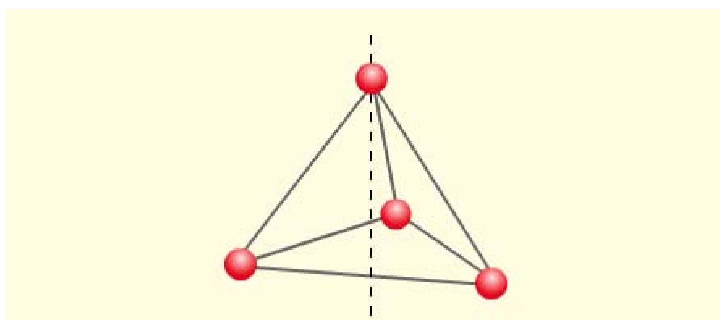
5.2 For the same arrangement as in the previous problem, what is the moment of inertia of the system of balls for the axis that is perpendicular to the plane of the square, and passes through its center point?

__________ kg · m^2

5.3 Four balls are connected by a straight rod. One end of the rod is painted blue. The first ball has mass 1.0 kg and is 1.0 m from the blue end of the rod, the second ball has mass 2.0 kg and is 2.0 m from the blue end, and so on for the other two balls. The mass of the rod is negligible compared to the mass of the balls. What is the moment of inertia of this system for an axis of rotation perpendicular to the rod and touching the blue end?

__________ kg · m^2

5.4 Four small balls of identical mass 2.36 kg are arranged in a rigid structure as a regular tetrahedron. (A regular tetrahedron has four faces, each of which is an equilateral triangle.) Each edge of the tetrahedron has length 3.20 m. What is the moment of inertia of the system, for an axis of rotation passing perpendicularly through the center of one of the faces of the tetrahedron?



__________ kg · m^2

5.5 A rigid rod of negligible mass runs through n balls, where the n^{th} ball has mass n kg. The balls are 1 m apart on the rod. Imagine that the rod extends 1 m beyond the 1 kg ball and that an axis of rotation passes through that end of the rod, perpendicular to it. Find a closed-form expression for the moment of inertia of the system as a function of n.

Section 6 - A table of moments of inertia

6.1 A tire of mass 1.3 kg and radius 0.34 m experiences a constant net torque of 2.1 N·m. Treat the tire as though all of its mass is concentrated at its rim. How long does it take for the tire to reach an angular speed of 18 rad/s from a standing stop?

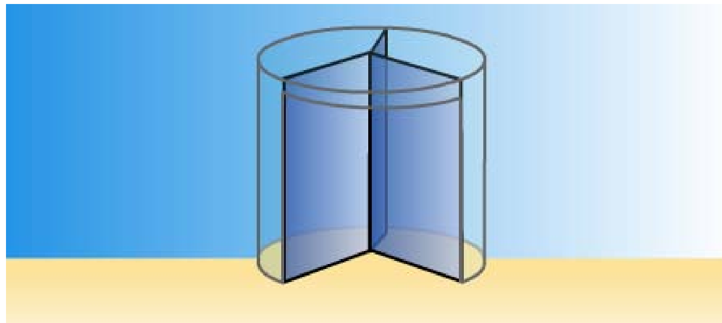
__________ s

6.2 A hollow glass holiday ornament in the shape of a sphere is suspended on a string that forms an axis of rotation through the ornament's center. If the ornament has mass 0.0074 kg and radius 0.036 m, what is its moment of inertia for an axis that passes through its center?

__________ kg · m^2

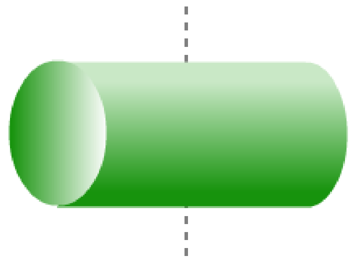
6.3 Three glass panels, each 2.7 m tall and 1.6 m wide and with mass 23 kg, are joined together to make a rotating door. The axis of rotation is the line where the panels join together. A person pushes perpendicular to a panel at its outer edge with a force of 73 N. What is the angular acceleration of the door?

__________ rad/s^2



6.4 A solid cylinder 0.230 m long and with a radius of 0.0730 m is rotated around an axis through its middle and parallel to its ends, as shown. A constant net torque of 1.20 N·m is applied to the cylinder, resulting in an angular acceleration of 23.0 rad/s^2. What is the mass of the cylinder?

__________ kg



6.5 Two golf balls are glued together and rotated about an axis through the point where they are joined. The axis is tangent to both golf balls. The mass of a golf ball is 0.046 kg and its radius is 0.021 m. What is the moment of inertia of the pair of golf balls around the chosen axis?

__________ $kg \cdot m^2$

6.6 Bob and Ray push an ordinary door from opposite sides. Both of them push perpendicular to the door. Bob pushes 0.670 m from the hinge with a force of 121 N. Ray pushes 0.582 m from the hinge with force of 132 N. Consider the door as a slab, with height 2.03 m, width 0.813 m, and mass 13.6 kg. What is the magnitude of the angular acceleration of the door?

__________ rad/s^2

6.7 A potter's wheel consists of a uniform concrete disk, 0.035 m thick and with a radius of 0.42 m. It has a mass of 48 kg. The wheel is rotating freely with an angular velocity of 9.6 rad/s when the potter stops it by pressing a wood block against the edge of the wheel, directing a force of 65 N on a line toward the center of the wheel. If the wheel stops in 7.2 seconds, what is the coefficient of friction between the block and the wheel?

____

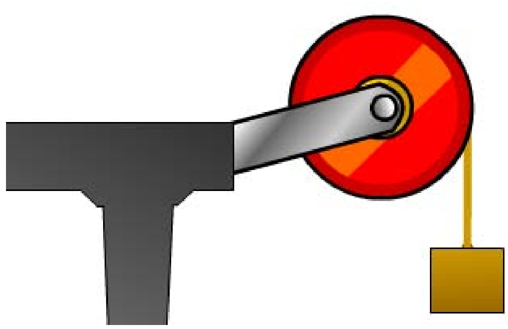
6.8 Two thin, square slabs of metal, each with side length of 0.30 m and mass 0.29 kg, are welded together in a T shape and rotated on an axis through their line of intersection. What is the moment of inertia of the T?

__________ $kg \cdot m^2$

Section 9 - Sample problem: an Atwood machine

9.1 The pulley in the illustration is a uniform disk of mass 2.8 kg and radius 0.24 m, which is free to rotate without friction. The mass of the block is 1.2 kg. Downward and clockwise are negative directions. (a) What is the angular acceleration of the pulley? (b) What is the acceleration of the falling block?

(a) __________ rad/s^2
(b) __________ m/s^2



Section 10 - Interactive problem: close the bridge

10.1 Use the simulation in the interactive problem in this section to answer the following question. What is the force required to close the bridge with the desired angular acceleration? Test your answer using the simulation.

__________ N

Section 11 - Parallel axis theorem

11.1 What is the formula for the moment of inertia of a hollow sphere of mass M and radius R, rotated on an axis tangent to its surface?

- ○ $(4/3)MR^2$
- ○ $(31/15)MR^2$
- ○ $(5/3)MR^2$
- ○ $(14/3)MR^2$
- ○ $2MR^2$

11.2 A thin rectangular slab, with dimensions 0.580 m by 0.830 m and mass 0.150 kg, is rotated about an axis passing through the slab parallel to the short edge. If the axis is 0.230 m from the short edge, what is the moment of inertia of the slab?

__________ kg · m^2

11.3 A solid ball of mass 0.68 kg and radius 0.17 m is attached by a small loop on its surface to a string of length 0.29 m. You are whirling the ball around in a circle by holding the end of the string.What is the moment of inertia of the ball? You can assume that the mass of the string and loop are negligible, but the sphere is large enough compared to the string length that it should not be considered as a particle.

__________ kg · m^2

11.4 Using the parallel axis theorem, prove that any axis of rotation that minimizes the moment of inertia for an object must pass through its center of mass.

11.5 A thin rod, 1.20 m long with mass 0.120 kg, is attached perpendicularly to the surface of a solid sphere of mass 0.390 kg and radius 0.180 m. The system is rotated around an axis perpendicular to the end of the rod opposite the sphere. What is the moment of inertia of the system?

__________ kg · m^2

Section 12 - Rotational work

12.1 A child pushes a merry-go-round with a force of 45.0 N at an angle tangent to the circle (that is, perpendicular to a radius). If the child pushes it through exactly one full circle, and the merry-go-round has a radius of 1.15 m, how much work does she do?

__________ J

12.2 In an automobile, the crankshaft transfers energy from the engine to the transmission, which in turn rotates the wheel axle. A particular car is rated at 126 hp (= 9.40×10^4 W), which you can take as the rate of energy transfer by the crankshaft when rotating at its maximum of 6050 revolutions per minute. What torque does the crankshaft supply?

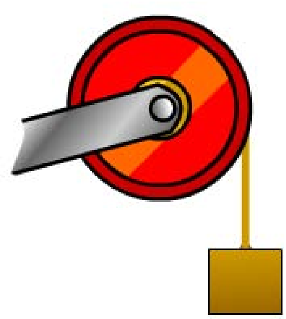
__________ N · m

12.3 A large rectangular bank vault door, 2.20 m wide, is hinged on one edge. To close the door, the bank manager applies a constant force of 161 N at the opposite edge of the door, perpendicular to the door. If the door moves through an angle of 135° as it closes, how much work does the bank manager do on the door?

__________ J

12.4 The pulley in the illustration is a uniform disk of mass 2.40 kg and radius 0.220 m. The block applies a contant torque to the pulley, which is free to rotate without friction, resulting in an angular acceleration of magnitude 0.180 rad/s^2 for the pulley. As the block falls 0.500 m, how much work does it do on the pulley?

__________ J



Section 13 - Rotational kinetic energy

13.1 A hollow closed cylinder of radius R is rolling without slipping. (Think of an empty tin can.) Each of the two endcaps has a mass of M, and the rest (the hollow tube) has a mass of $3M$, for a total object mass of $5M$. What is the ratio of linear kinetic energy to rotational kinetic energy? Enter your answer as a decimal number.

____

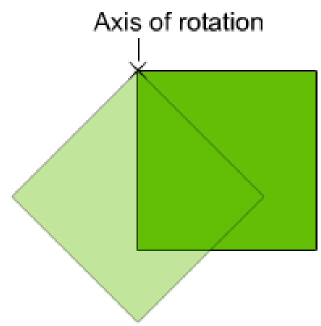
13.2 A certain pulley is a uniform disk of mass 2.7 kg and radius 0.25 m. A rope applies a constant torque to the pulley, which is free to rotate without friction, resulting in an angular acceleration of 0.12 rad/s^2. The pulley starts at rest at time t = 0 s. What is its rotational kinetic energy at t = 2.1 s?

__________ J

13.3 You hold a rod, 1.20 m long, in a vertical position with its bottom end touching a carpeted floor. You then release it so it falls over, which it does without the bottom end slipping. What is the magnitude of the angular velocity of the thin rod when it hits the carpet? (Use the principle of conservation of energy.)

__________ rad/s

13.4 A thin square slab with sides of length 0.550 m and mass 1.80 kg is suspended so it can freely rotate about a horizontal axis of rotation perpendicular to the surface at one corner. The square is held so its top and bottom edges are horizontal, and then is released. What is the angular velocity of the square when the point opposite the pivot is as low as possible?



__________ rad/s

13.5 A 0.870 kg object is attached to a rope that is wrapped around a frictionless pulley. The pulley is a solid 0.450 kg disk with a 0.140 meter radius. If the object is held at rest, and then released, what will its speed be after it falls 0.120 meters?

__________ m/s

13.6 A thin string is wrapped several times around a pulley and is then attached to a block of mass 3.00 kg on one side. The pulley is a uniform disk of mass 15.0 kg and diameter 0.700 m, and turns on a frictionless pivot. The block is released from rest, falls a height h, and hits the floor. Just before the block hits, it is moving at 2.65 m/s. What is the height h?

__________ m

Section 15 - Rolling objects and kinetic energy

15.1 A 5.75 kg solid ball with a radius of 0.185 m rolls without slipping at 3.55 m/s. What is its total kinetic energy?

__________ J

15.2 A thin hoop of mass 7.8 kg rolls on a horizontal floor with a speed (at its center of mass) of 0.23 m/s. How much work must be done on the hoop to bring it to a stop? Reminder: Work done on an object that increases the object's *KE* is positive and work that decreases the *KE* is negative.

__________ J

15.3 A solid ball, with radius r = 6.00 centimeters and mass m = 0.400 kg, rolls without slipping on the inside of a fixed horizontal circular pipe with radius R = 0.300 meters. The ball is held at rest with its center lying on the horizontal diameter of the pipe. It is then released and rolls down the side of the pipe, in a plane perpendicular to the length of the pipe. (a) What is the total kinetic energy of the ball when it reaches the bottom of the pipe? Hint: You must take into account the distance between the ball's center of mass and the inner surface of the pipe. (b) What is the speed of the ball's center of mass when it reaches the bottom? (c) What normal force does the pipe exert on the ball when the ball reaches the bottom?

(a) __________ J
(b) __________ m/s
(c) __________ N

Section 19 - Physics at play: a yo-yo

19.1 A strange yo-yo is made up of two identical solid spheres connected by a short spindle. The radius of the spheres is 0.0340 m and each of them has mass 0.00750 kg. The spindle has radius 0.00550 m and mass 0.00350 kg, and the string is wrapped around the spindle. What is the magnitude of the acceleration of the yo-yo as it unreels and rolls down the string?

__________ m/s^2

Section 21 - Angular momentum of a particle in circular motion

21.1 What is the magnitude of angular momentum of a 1070 kg car going around a circular curve with a 15.0 m radius at 12.0 m/s? Assume the origin is at the center of the curve's arc.

__________ kg · m^2 /s

21.2 Calculate the magnitude of the angular momentum of the Earth around the Sun, using the Sun as the origin. The Earth's mass is 5.97×10^{24} kg and its roughly circular orbit has a radius of 1.50×10^{11} m. Use a 365-day year with exactly 24 hours in each day.

_________ $kg \cdot m^2/s$

Section 22 - Angular momentum of a rigid body

22.1 A thin rod 2.60 m long with mass 3.80 kg is rotated counterclockwise about an axis through its midpoint. It completes 3.70 revolutions every second. What is the magnitude of its angular momentum?

_________ $kg \cdot m^2/s$

22.2 A 0.16 meter long, 0.15 kg thin rigid rod has a small 0.22 kg mass stuck on one of its ends and a small 0.080 kg mass stuck on the other end. The rod rotates at 1.7 rad/s through its physical center without friction. What is the magnitude of the angular momentum of the system taking the center of the rod as the origin? Treat the masses on the ends as point masses.

_________ $kg \cdot m^2/s$

22.3 A solid uniform cylinder rolls down a ramp, starting from a stationary position at height 1.4 m. At the bottom of the ramp, the cylinder has no potential energy. If the mass of the cylinder is 2.9 kg and its radius is 0.076 m, what is the magnitude of angular momentum of the cylinder at the bottom of the ramp with respect to the cylinder's center of mass?

_________ $kg \cdot m^2/s$

22.4 A solid cylinder is rotated counterclockwise around an axis at its base (one circular end). The axis lies in the plane of the base, and passes through the center of the circle. The cylinder is 3.60 m long and has a radius of 0.870 m. Its mass is 2.30×10^4 kg and the center of the base opposite the axis has a tangential speed of 12.0 m/s. What is the cylinder's angular momentum?

_________ $kg \cdot m^2/s$

Section 24 - Sample problem: object moving in a straight line

24.1 Two particles move along parallel straight-line paths, but in opposite directions. Particle A has mass 0.220 kg and constant velocity 1.40 m/s. Particle B has mass 0.120 kg and constant velocity of magnitude 1.80 m/s. Their paths are 1.30 m apart. What is the magnitude of the angular momentum of the system with respect to a point midway between the two parallel lines and in the same plane?

_________ $kg \cdot m^2/s$

24.2 You hold a golf ball of mass 0.046 kg at arm's length straight out, and drop it. After 0.75 seconds, what is the magnitude of the ball's angular momentum with respect to (a) your hand and (b) your shoulder, which is 0.65 m from your hand and at the same height?

(a) _________ $kg \cdot m^2/s$
(b) _________ $kg \cdot m^2/s$

Section 26 - Torque and angular momentum

26.1 An electric drill delivers a net torque of 15.0 N·m to a buffing wheel used to polish a car. The buffing wheel has a moment of inertia of 2.30×10^{-3} kg·m². At 0.0220 s after the drill is turned on, what is the angular velocity of the buffing wheel?

_________ rad/s

26.2 A merry-go-round is 18.0 m in diameter and has a mass, unloaded, of 48,100 kg. It is fairly uniform in structure and can be considered to be a solid cylindrical disk. The merry-go-round carries 28 passengers, who have an average mass of 68.5 kg and all sit at a distance of 8.75 m from the center. The fully-loaded merry-go-round takes 48.0 s to reach an angular velocity of 0.650 rad/s. (a) What is the moment of inertia of the loaded merry-go-round? (b) What constant net torque is applied to the merry-go-round to reach this working speed?

(a) _________ $kg \cdot m^2$
(b) _________ $N \cdot m$

26.3 A string is wound around the edge of a solid 1.60 kg disk with a 0.130 m radius. The disk is initially at rest when the string is pulled, applying a force of 6.50 N in the plane of the disk and tangent to its edge. If the force is applied for 1.90 seconds, what is the magnitude of its final angular velocity?

_________ rad/s

26.4 The platter of an old-fashioned turntable for playing records is driven by a small rotating wheel, called a drive wheel, at its circumference. The platter is 0.30 m in diameter and has mass 0.32 kg. If the platter starts from rest and rotates friction-free and takes 0.58 s to reach its full speed of 33 1/3 rpm, what is the magnitude of the force applied by the drive wheel?

_________ N

Section 27 - Conservation of angular momentum

27.1 A 1.6 kg disk with radius 0.63 m is rotating freely at 55 rad/s around an axis perpendicular to its center. A second disk that is not rotating is dropped onto the first disk so that their centers align, and they stick together. The mass of the second disk is 0.45 kg and its radius is 0.38 m. What is the angular velocity of the two disks combined?

__________ rad/s

27.2 Two astronauts in deep space are connected by a 22 m rope, and rotate at an angular velocity of 0.48 rad/s around their center of mass. The mass of each astronaut, including spacesuit, is 97 kg, and the rope has negligible mass. One astronaut pulls on the rope, shortening it to 14 m. What is the resulting angular velocity of the astronauts?

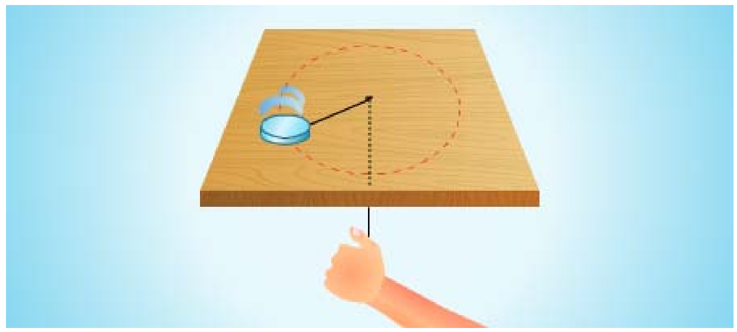
__________ rad/s

27.3 A 27.5 kg child stands at the center of a 125 kg playground merry-go-round which rotates at 3.10 rad/s. If the child moves to the edge of the merry-go-round, what is the new angular velocity of the system? Model the merry-go-round as a solid disk.

__________ rad/s

27.4 A puck with mass 0.28 kg moves in a circle at the end of a string on a frictionless table, with radius 0.75 m. The string goes through a hole in the table, and you hold the other end of the string. The puck is rotating at an angular velocity of 18 rad/s when you pull the string to reduce the radius of the puck's travel to 0.55 m. Consider the puck to be a point mass. What is the new angular velocity of the puck?

__________ rad/s



Section 29 - Interactive summary problem: dynamics of skating

29.1 Using the information given in the first interactive problem in this section, what is the correct distance for your final arm position to attain the correct angular velocity? Test your answer using the simulation.

__________ m

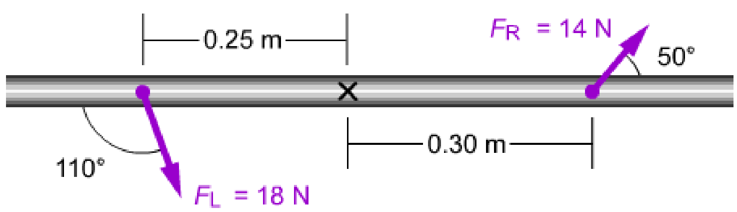
29.2 Using the information given in the second interactive problem in this section, what is the frictional force needed to come to a complete stop in the required number of revolutions? Test your answer using the simulation.

__________ N

Additional Problems

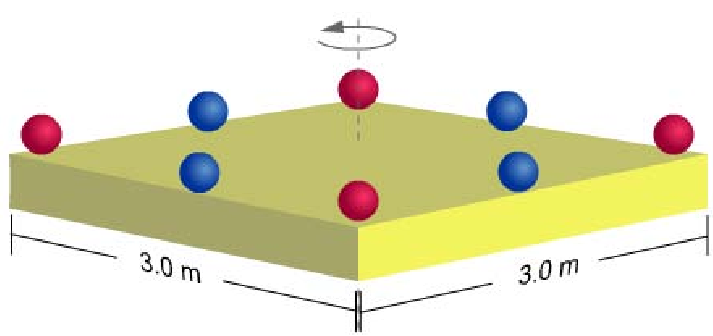
A.1 If the angular acceleration of the rod pictured is 3.7 rad/s^2, what is its moment of inertia?

__________ kg · m^2



A.2 A 20 kg square slab with side lengths 3.0 m rotates around an axis perpendicular to the slab. A particle of mass 0.60 kg sits on each of the corners of the square. Also, a particle of mass 0.80 kg sits at the center of each of the edges. What is the moment of inertia of the system?

__________ kg · m^2



A.3 A 0.555 kg hoop of radius 0.640 m is spinning in a vertical plane (that is, around a horizontal axis) with an angular speed ω_o = 20.0 rad/s and then drops a very short distance straight down onto a rough, level floor. The hoop slips for some time, starts picking up speed in the horizontal direction, and then eventually rolls without slipping. (a) Calculate the final speed of its center of mass, v_f. (b) How much work does friction perform during the process? Ignore the effects of air resistance.

(a) __________ m/s
(b) __________ J

Chapter 12 Static Equilibrium and Elasticity

12.0 - Introduction

A tightrope walker balances forces and torques to maintain equilibrium.

Although much of physics focuses on motion and change, the topic of how things stay the same − equilibrium − also merits study. Bridges spanning rivers, skyscrapers standing tall… None of these would be possible without engineers having achieved a keen understanding of the conditions required for equilibrium. In this chapter, we focus on static equilibrium. To do so, we must consider both forces and torques, since for an object to be in static equilibrium, the net force and net torque on it must both equal zero.

The forces and masses involved in equilibrium can be stupendous. The Brooklyn Bridge was the engineering marvel of its day; it gracefully spans the East River between Brooklyn and Manhattan. The anchorages at the ends of the bridge each have a mass of almost 55 million kilograms, while the suspended superstructure between the anchorages has a mass of 6 million kilograms. Supporting those 6 million kilograms are four cables of 787,000 kilograms apiece. A five-year construction effort resulted in the largest suspension bridge of its time, and one that over 200,000 vehicles pass over daily.

Engineers who design structures such as bridges must concern themselves with forces that cause even a material like steel to change shape, to lengthen or contract. If the material returns to its initial dimensions when the force is removed, it is called elastic.

12.1 - Static equilibrium

Static equilibrium: No net torque, no net force and no motion.

In California, equilibrium is achieved either by renouncing one's possessions, moving to a commune and selecting a guru, or by becoming extremely rich, moving to Malibu and choosing a personal trainer.

In physics, static equilibrium also requires a threefold path. First, there is no net force acting on the body. Second, there is no net torque on it about any axis of rotation. Finally, in the case of static equilibrium, there is no motion. An object moving with a constant linear and rotational velocity is also in equilibrium, but not in *static* equilibrium.

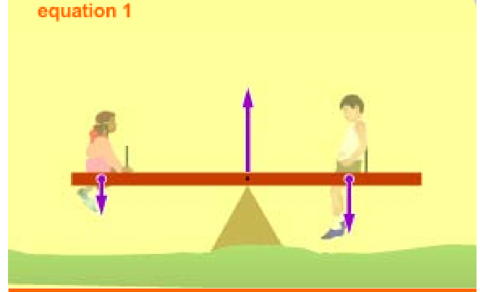
Let's see how we can apply these concepts to the seesaw at the right. There are two children of different weights on the seesaw. They have adjusted their positions so that the seesaw is stationary in the position you now see. (We will only concern ourselves with the weights of the children, and will ignore the weight of the seesaw.)

In Equation 2, we examine the torques. The fulcrum is the axis of rotation. Since the system is stationary, there is no angular acceleration, which means there is no net torque.

Let's consider the torques in more detail. They must sum to zero since the net torque equals zero. We choose to use an axis of rotation that passes through the point where the fulcrum touches the seesaw. The normal force of the fulcrum creates no torque because its distance to this axis is zero. The boy exerts a clockwise (negative) torque. The girl exerts a counterclockwise (positive) torque. Since the girl weighs less than the boy, she sits farther from the fulcrum to make their torques equal but opposite. In sum, there are no net forces, no net torques, and the system is not moving: It is in static equilibrium.

Note that in analyzing the seesaw, we used the fulcrum as the axis of rotation. This seems natural, since it is the point about which the seesaw rotates when the children are "seesawing." However, in problems you will encounter later, it is not always so easy to determine the axis of rotation. In those cases, it is helpful to remember that if the net torque is zero about one axis, it will be zero about any axis, so the choice of axis is up to you. However, this trick only works for cases when the net torque is zero. In general, the torque depends on one's choice of axis.

equation 1

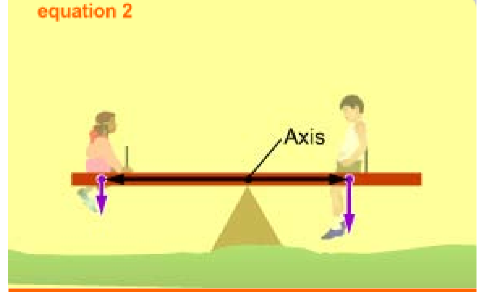


Static equilibrium

$$\Sigma F_x = 0, \Sigma F_y = 0$$

Net force along each axis is zero

equation 2

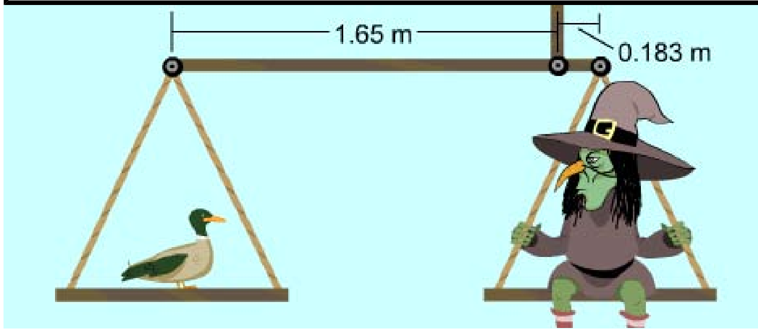


Static equilibrium

$$\Sigma\tau = 0$$

Net torque is zero

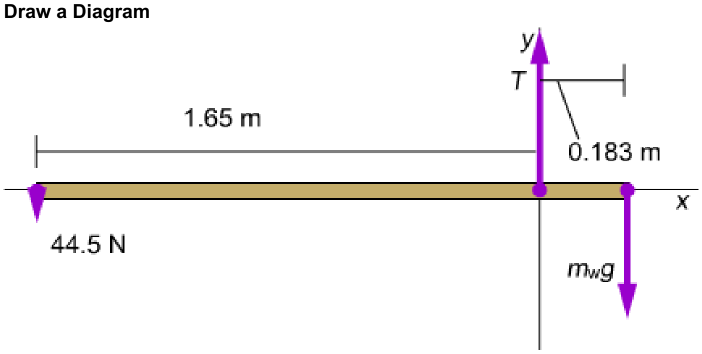
12.2 - Sample problem: a witch and a duck balance



The witch and the duck are balanced on the scale. The duck weighs 44.5 N. What is the witch's weight?

Since the system is stationary, it is in static equilibrium. This means there is no net torque and no net force. We use the pivot point of the scale as the axis of rotation. We start by drawing a diagram for the problem. (The diagrams in this section are not drawn to scale; we also are not really sure witches exist, but then we do assume objects to be massless and frictionless and so on, so who are we to complain?) We ignore the masses of the various parts of the balance in this problem.

Draw a Diagram



Variables

Sign is important with torques: the duck's torque is in the positive (counterclockwise) direction while the witch's torque is in the negative (clockwise) direction. To calculate the magnitude of each torque, we can multiply the force by the lever arm, since the two are perpendicular.

Forces	x	y
weight, duck	0 N	$-m_dg = -44.5$ N
weight, witch	0 N	$-m_wg$
tension	0 N	T
Torques	lever arm (m)	torque (N·m)
weight, duck	1.65	(1.65)(44.5)
weight, witch	0.183	$-(0.183)(m_wg)$
tension	0	0

What is the strategy?

1. Draw a free-body diagram that shows the forces on the balance beam.
2. Place the axis of rotation at the location of an unknown force (the tension). This simplifies solving the problem. There is no need to calculate the amount of this force since a force applied at the axis of rotation does not create a torque.
3. Use the fact that there is no net torque to solve the problem. The only unknown in this equation is the weight of the witch.

Physics Principles and Equations

There is no net torque since there is no angular acceleration.

$$\Sigma\tau = 0$$

The weights are perpendicular to the beam, so we calculate the torques they create using

$$\tau = rF$$

A force (like tension) applied at the axis of rotation creates no torque.

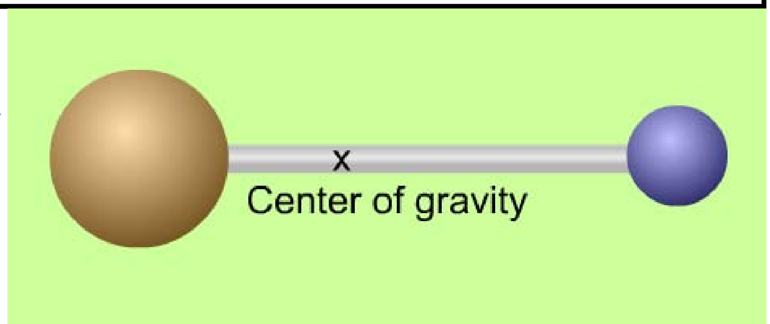
Step-by-step solution

We use the fact that the torques sum to zero to solve this problem. There are only two torques in the problem due to the location of the axis of rotation: The tension creates no torque.

Step	Reason
1. torque of witch + torque of duck = 0	no net torque
2. $-(0.183\text{ m})(m_w g) + (1.65\text{ m})(44.5\text{ N}) = 0\text{ N}\cdot\text{m}$	substitute values
3. $m_w g = (1.65\text{ m})(44.5\text{ N}) / (0.183\text{ m})$	solve for $m_w g$
4. $m_w g = 401\text{ N}$	evaluate

12.3 - Center of gravity

Center of gravity: The force of gravity effectively acts at a single point of an object called the center of gravity.



The center of gravity of a barbell.

The concept of center of gravity complements the concept of center of mass. When working with torque and equilibrium problems, the concept of center of gravity is highly useful.

Consider the barbell shown above. The sphere on the left is heavier than the one on the right. Because the spheres are not equal in weight, if you hold the barbell exactly in the center, the force of gravity will create a torque that causes the barbell to rotate. If you hold it at its center of gravity however, which is closer to the left ball than to the right, there will be no net torque and no rotation.

When a body is symmetric and uniform, you can calculate its center of gravity by locating its geometric center. Let's consider the barbell for a moment as three distinct objects: the two balls and the bar. Because each of the balls on the barbell is a uniform sphere, the geometric center of each coincides with its center of gravity. Similarly, the center of gravity of the bar connecting the two spheres is at its midpoint.

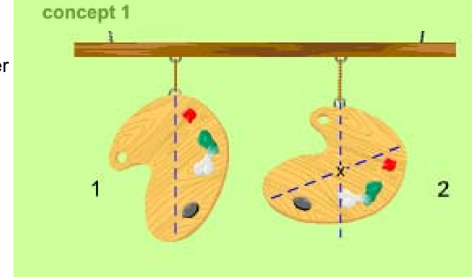
When we consider the entire barbell, however, the situation gets more complicated. To calculate the center of gravity of this entire system, you use the equation to the right. This equation applies for any group of masses distributed along a straight line. To apply the equation, pick any point (typically, at one end of the line) as the origin and measure the distance to each mass from that point. (With a symmetric, uniform object like a ball, you measure from the origin to its geometric center.)

Then, multiply each distance by the corresponding weight, add the results, and divide that sum by the sum of the weights. The result is the distance from the origin you selected to the center of gravity of the system. The center of gravity of an object does not have to be within the mass of the object: For example, the center of gravity of a doughnut is in its hole.

If you have studied the center of mass, you may think the two concepts seem equivalent. They are. When g is constant across an object, its center of mass is the same as its center of gravity. Unless the object is enormous (or near a black hole where the force of gravity changes greatly with location), a constant g is a good assumption.

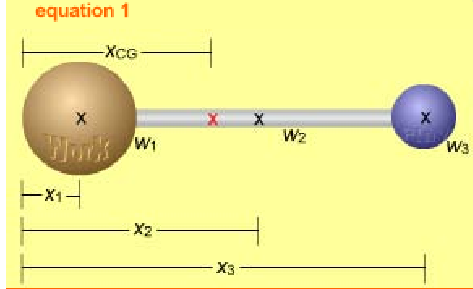
You can empirically determine the center of gravity of any object by dangling it. In Concept 1, you see the center of gravity of a painter's palette being determined by dangling. To find the center of gravity of an object using this method, hang (dangle) the object from a point and allow it to move until it naturally stops and rests in a state of equilibrium. The center of gravity lies directly below the point where the object is suspended, so you can draw an imaginary line through the object straight down from the point of suspension. The object is then dangled again, and you draw another line down from the suspension point. Since both of these lines go through the center of gravity, the center of gravity is the point where the lines intersect.

concept 1



Center of gravity
One point where weight effectively acts
Can be found by "dangling" object twice

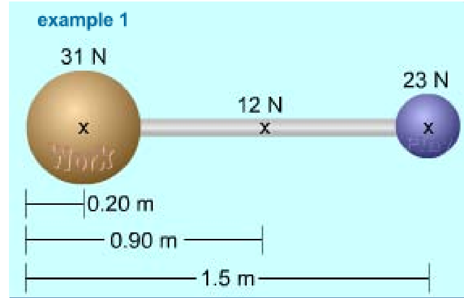
equation 1



Center of gravity

$$x_{CG} = \frac{w_1x_1 + w_2x_2 + \ldots + w_nx_n}{w_1 + w_2 + \ldots + w_n}$$

x_{CG} = x position of center of gravity
w_i = weight of object i
x_i = x position of object i



Where is the center of gravity?

$$x_{CG} = \frac{w_1x_1 + w_2x_2 + ... + w_nx_n}{w_1 + w_2 + ... + w_n}$$

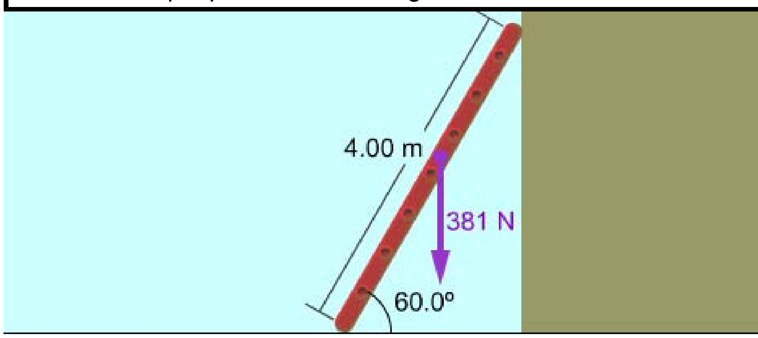
Put the origin at center of gold ball:

$$x_{CG} = \frac{(31\ \text{N})(0\ \text{m}) + (12\ \text{N})(0.70\ \text{m}) + (23\ \text{N})(1.3\ \text{m})}{31 + 12 + 23\ \text{N}}$$

$x_{CG} = (8.4 + 29.9) / 66$

$x_{CG} = 0.58$ m from center of gold-colored ball

12.4 - Sample problem: a leaning ladder



There is no friction between the wall and the ladder. What are the components of the force of the ground on the ladder?

In this example, the ladder leans against a smooth (that is, frictionless) wall and is not moving. “Frictionless” means the wall exerts no vertical force on the ladder. However, it does exert a horizontal force, a normal force on the ladder.

The diagram shows the weight of the ladder and the angle it makes with the ground. Assume the ladder to be uniform. Its center of gravity is at its midpoint. You are asked to find the components of the force that the ground exerts on the ladder. The x component of ground force is a static friction force; the y component is a normal force.

This is a classic physics problem. Here are some tips and techniques that prove useful in general.

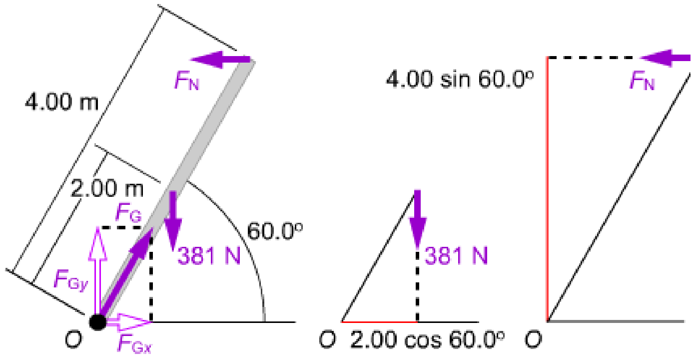
First: State equations for the forces acting in each dimension. The horizontal forces sum to zero, and the vertical forces sum to zero. These are your first two equations.

Second: Choose your axis of rotation carefully. You can choose any one you like since the system is in equilibrium and is not rotating. The “trick” is to choose an axis at a point where unknown forces act, since they will exert no torque (forces acting at the axis of rotation do not exert a torque). For instance, in this problem, the point where the ladder touches the ground is chosen, since it means we can ignore the two forces that act there. Those forces are unknown, so being able to ignore them simplifies things. Choosing this axis means there is only one unknown force, and you can solve for it.

Third: Using the chosen axis of rotation, sum the torques to zero to write a third equation.

There are problems that are easier and harder to solve than this one, but these are some general techniques that often prove useful.

Draw a diagram



Above on the left is the free-body diagram, and on the right are two diagrams to help you find the lever arms for the weight and normal force vectors. For the sake of visual clarity, we have not drawn the x and y axes.

We chose the point where the ladder meets the ground for the axis of rotation. This means the ground force creates no torque. We can choose this axis because there is no net torque, so the choice of axis is irrelevant.

Variables

Forces	x	y
weight, ladder	0	$-mg = -381$ N
normal force from wall	$-F_N$	0
ground force	F_{Gx} (friction)	F_{Gy} (normal)
Torques	lever arm (m)	torque (N·m)
weight, ladder	2.00 cos 60.0°	−(381)(2.00 cos 60.0°)
normal force	4.00 sin 60.0°	F_N(4.00 sin 60.0°)
ground force	0	0

What is the strategy?

1. Draw a free-body diagram showing the forces on the ladder. Since the situation is somewhat complicated, draw separate diagrams to find the lever arms for the torques.
2. Since this is a static equilibrium situation, the net force in the x direction and in the y direction both equal zero. This gives you two equations to use.
3. State an equation for torque, using the fact that the net torque also equals zero. We choose the point where the ladder touches the ground as our axis of rotation. Choosing this point simplifies the torque calculations since we can ignore both forces acting there, a good thing, since we know neither of them.
4. To calculate the normal force of the ground, we consider forces acting vertically. Because there is just one unknown in this equation, it can be solved.
5. There are two unknown forces, so we have to write two equations in order to determine these forces.

Physics principles and equations

In equilibrium, the forces sum to zero in the x and y directions

$$\Sigma F_x = 0, \Sigma F_y = 0$$

In equilibrium the torques sum to zero

$$\Sigma\tau = 0$$

Step-by-step solution

The forces in the y direction sum to zero, which lets us quickly find the y component of the ground force (the normal force from the ground).

Step	Reason
1. $-mg + F_{Gy} = 0$	equilibrium applied in y direction
2. $F_{Gy} = 381$ N	solve for normal ground force

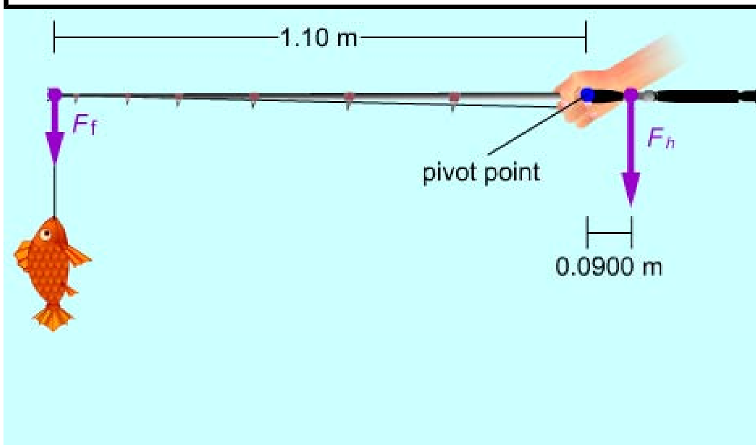
To find the x component of the ground force, the friction force, we first use the fact that all the forces sum to zero in the x direction. This provides one equation for the x component, but with two unknowns − we still need another equation.

Step	Reason
3. $-F_N + F_{Gx} = 0$	equilibrium applied in x direction
4. $F_{Gx} = F_N$	solve for frictional ground force

We use the fact that there is no net torque to get a second equation involving the x component of the ground force. The torques are created by the ladder's weight and the normal force from the wall. Using both equations, we find the actual value of the x component of the ground force.

Step	Reason
5. $\Sigma\tau = 0$	no net torque
6. $F_N(4.00 \sin 60.0°) - (381)(2.00 \cos 60.0°) = 0$	substitute values
7. $F_{Gx}(4.00 \sin 60.0°) - (381)(2.00 \cos 60.0°) = 0$	substitute F_{Gx} for F_N, from step 4
8. $F_{Gx} = \dfrac{(381)(2.00 \cos 60.0°)}{4.00 \sin 60.0°}$	solve for F_{Gx}
9. $F_{Gx} = 110$ N	evaluate

12.5 - Interactive checkpoint: fishing



Congratulations, you caught a 5.00 kg fish! When you hold the fish so that it hangs at the end of your relatively light fishing rod, you create a pivot point with your hand at the point shown on the diagram. You also exert a force on the rod with the heel of your hand where shown. If the fish hangs 1.10 m from the pivot point, and there are 0.0900 m between the pivot and the point at which the heel of your hand applies force, how much force do you have to exert in order for the rod to remain horizontal?

Answer:

F = ________ N

12.6 - Sample problem: isometric training

The lower arm and dumbbell are not moving. What amount of force does the biceps muscle exert?

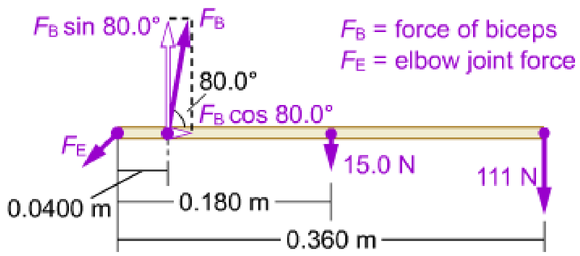
A weightlifter holds a stationary barbell. The lower arm (elbow to hand) is horizontal and pivots at the elbow. The biceps muscle applies a force at an angle of $80.0°$, 0.0400 m from the elbow. You are asked to find the amount of force supplied by the biceps.

The lower arm is roughly uniform in structure, so assume its weight is uniformly distributed and use its midpoint for its center of gravity when calculating torque.

To solve the problem, we use the elbow as the axis of rotation. We can do so because when there is no net torque, we can choose the axis of rotation.

With the axis of rotation at the elbow joint, the torque created by the elbow force is zero. The x component of the biceps force also creates no torque because it has no lever arm.

Diagram



To simplify the drawing, we did not draw the x and y axes.

Variables

Forces	x	y
weight, lower arm	0	−15.0 N
weight, dumbbell	0	−111 N
biceps force	$F_B \cos 80.0°$	$F_B \sin 80.0°$
elbow force	$-F_{Ex}$	$-F_{Ey}$

Torques	lever arm (m)	torque (N·m)
weight, lower arm	0.180	−(15.0)(0.180)
weight, dumbbell	0.360	−(111)(0.360)
biceps force, y	0.0400	$(F_B \sin 80.0°)(0.0400)$
elbow force	0	0

What is the strategy?

1. Draw a free-body diagram that shows the forces on the lower arm.
2. Since the system is in static equilibrium, the net forces in the x direction and in the y direction both equal zero. This gives you two possible equations to use.
3. For the same reason, the net torque also equals zero. Choose the elbow joint for the axis of rotation, which simplifies the calculations. The elbow force and the x component of the biceps force do not create any torque for this axis of rotation.
4. Look at the equations. Notice that, using the torques, you can write an equation that has only one unknown.

Physics principles and equations

The torques sum to zero.

$$\Sigma\tau = 0$$

The torque is a cross product. When the force and the displacement vector are parallel, there is no torque.

Step-by-step solution

Step	Reason
1. biceps torque + lower arm torque + dumbbell torque = 0	no net torque
2. $(F_B \sin 80.0°)(0.0400) - (15.0)(0.180) - (111)(0.360) = 0$	substitute values
3. $F_B = \dfrac{(15.0)(0.180) + (111)(0.360)}{(\sin 80.0°)(0.0400)}$	solve for force of biceps
4. $F_B = 1080$ N	evaluate

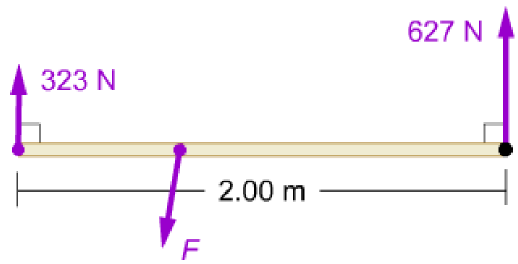
The force exerted by the biceps is approximately ten times the weight of the barbell.

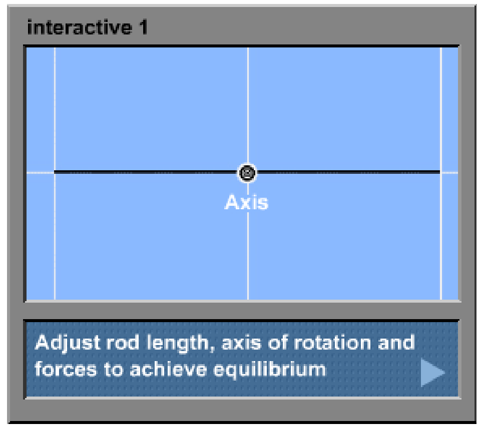
12.7 - Interactive problem: achieve equilibrium

In the simulation on the right, you are asked to apply three forces to a rod so that it will be in static equilibrium. Two of the forces are given to you and you have to calculate the magnitude, position, and direction of the third force. If you do this correctly, when you press GO, the rod will not move.

The rod is 2.00 meters long, and is horizontal. A force of 323 N is applied to the left end, straight up. A force of 627 N is applied to the right end, also straight up. You are asked to apply a force to the rod that will balance these two forces and keep it in static equilibrium.

Here is a free-body diagram of the situation. We have **not** drawn the third force where it should be!





After you calculate the third vector's magnitude, position and direction, follow these steps to set up the simulation.

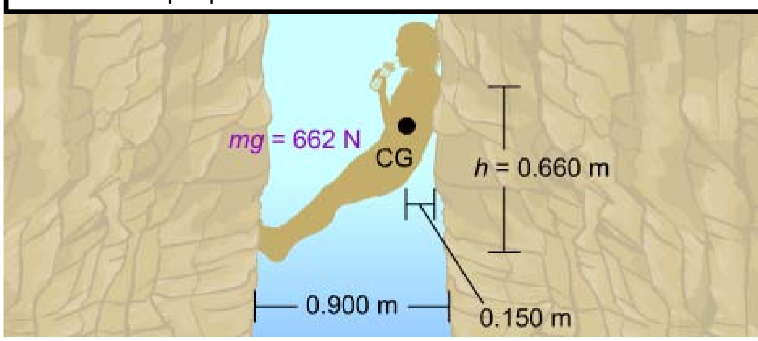
1. Adjust the rod length so it is 2.00 m.
2. Drag the axis of rotation to an appropriate position.
3. Apply all the forces. Drag a force vector by its tail from the control panel and attach the tail to the rod. You can then move the tail of the vector along the rod to the correct position, and drag the head of the vector to change its length and angle.

The control panel will show you the force's magnitude, direction and distance to the axis of rotation. The vector whose properties are being displayed has its head in blue.

When you have the simulation set up, press GO. If everything is set up correctly, the rod will be in equilibrium and will not move. Press RESET if you need to make any adjustments. If you have trouble, refer to the section on static equilibrium in this chapter, and the section on torque in the Rotational Dynamics chapter.

After you solve this interactive problem, consider the following additional challenge. What do you think will happen in the simulation if you change the position of the axis of rotation? Make a guess, and test your hypothesis with the simulation.

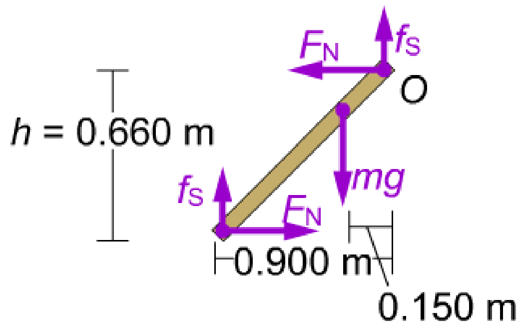
12.8 - Sample problem: a rock climber



The friction forces at the climber's feet and shoulders are equal. She presses with minimal force to stay in equilibrium. How much force does she exert on a rock wall?

A rock climber pauses in what climbers call a chimney by pressing her shoulders and feet on opposing rock walls. She presses with the minimum force needed to stay in static equilibrium; you are asked to find the amount of this force. This is the same as the amount of the normal force the walls exert on her.

Diagram



We indicate the equal friction forces at her feet and shoulders as f_s and the equal normal forces as F_N. To simplify the drawing, we do not draw the x and y axes. We place the axis of rotation at the climber's shoulders, which means the two forces acting there create no torque.

Variables

Forces	x	y
friction at feet	0	f_s
friction at shoulders	0	f_s
weight	0	$-mg = -662$ N
normal force at feet	F_N	0
normal force at shoulder	$-F_N$	0

Torques	lever arm (m)	torque (N·m)
friction at feet	0.900	$-f_s(0.900)$
friction at shoulders	0	0
weight	0.150	(662)(0.150)
normal force at feet	0.660	$F_N(0.660)$
normal force at shoulder	0	0

What is the strategy?

1. Draw a free-body diagram showing the forces on the climber.
2. Since she is in static equilibrium, the net force in the x direction and the net force in the y direction both equal zero. This gives you two possible equations to use.
3. The net torque also equals zero. This gives another possible equation. As mentioned, we place the axis of rotation at the climber's shoulders, which means the friction and normal forces there create no torque.
4. Consider the resulting equations to decide which are the most helpful. In this case, first find the friction forces using the net force equation for the y direction, since it contains only one unknown value.
5. Then use the torque equation to find the normal force. Since the static friction force was just calculated, there is now only one unknown in this equation.

Step-by-step solution

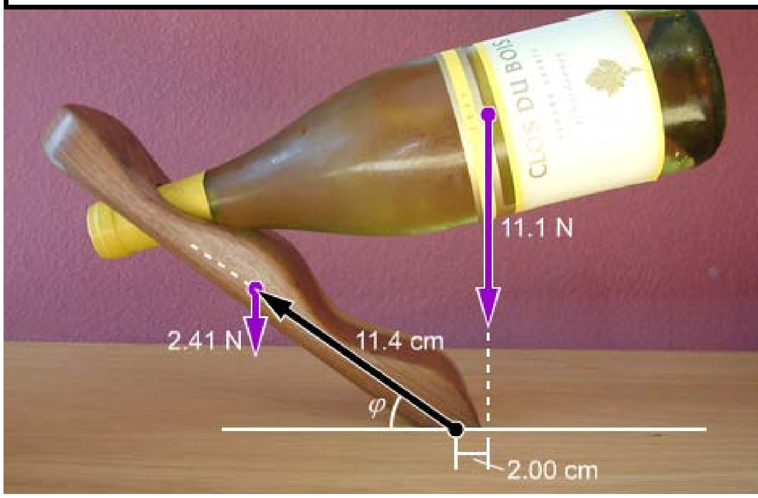
Forces are in equilibrium along the y axis. We use this fact to find the friction force.

Step	Reason
1. $f_s + f_s - mg = 0$	equilibrium applied in y direction
2. $2f_s - 662\text{ N} = 0$	substitute value
3. $f_s = 331\text{ N}$	solve for friction force

Now we use equilibrium for torque to find the normal force. We placed the axis of rotation at the climber's shoulder, so we can skip forces applied at that point when calculating torques.

Step	Reason
4. normal force torque + weight torque + friction torque = 0	no net torque
5. $F_N(0.660) + (662)(0.150) - f_s(0.900) = 0$	substitute torques
6. $F_N(0.660) + (662)(0.150) - (331)(0.900) = 0$	substitute equation 3 into equation 5
7. $F_N = \frac{(331)(0.900) - (662)(0.150)}{0.660\text{ m}}$	solve for normal force
8. $F_N = 301\text{ N}$	evaluate

12.9 - Interactive checkpoint: wine holder



The clever wine bottle holder pictured is in static equilibrium. Consider the axis of rotation to be at the center of the holder's base. The wine bottle weighs 11.1 N while the holder weighs 2.41 N. The wine bottle's center of gravity is a horizontal distance of 2.00 cm from the axis of rotation. The holder's center of gravity is 11.4 cm from the base, along the holder. What is the angle φ?

In order to keep a good seal, wine bottles need to be stored so that their corks remain wet. This holder stores a bottle in the required position, and in a manner that may cause you to do a double take!

Answer:

φ = ____ °

12.10 - Elasticity

Elastic: An elastic object returns to its original dimensions when a deforming force is removed.

concept 1

Elastic
Shape changes under force
Elastic: when force is removed, original shape is restored

In much of physics, the object being analyzed is assumed to be rigid and its dimensions are unchanging. For instance, if a homework problems asks: "What is the effect of a net force on a car?" and you answer, "The net force on the car caused a dent in its fender and enraged the owner," then you are thinking a little too much outside the box. The question anticipates that you will apply Newton's second law, not someone's insurance policy.

However, objects do stretch or compress when an external force is applied to them. They may return to their original dimensions when that force is removed: Objects with this property are called elastic. The external force causes the bonds between the molecules that make up the material to stretch or compress, and when that force is removed, the molecules can "spring back" to their initial configuration. The extent to which the dimensions of an object change in response to a *deforming force* is a function of the object's original dimensions, the material that makes it up, and the nature of the force that is applied.

It is also possible to stretch or compress an object to the point where it is unable to spring back. When this happens, the object is said to be *deformed*.

It can require a great deal of force to cause significant stretching. For instance, if you hang a 2000 kg object, like a midsize car, at the end of a two-meter long steel bar with a radius of 0.1 meters, the bar will stretch only about 6×10^{-6} meters.

There are various ways to change the dimensions of an object. For instance, it can be stretched or compressed along a line, like a vertical steel column supporting an overhead weight. Or, an object might experience compressive forces from all dimensions, like a ball submerged in water. Calculating the change of dimensions is a slightly different exercise in each case.

12.11 - Stress and strain

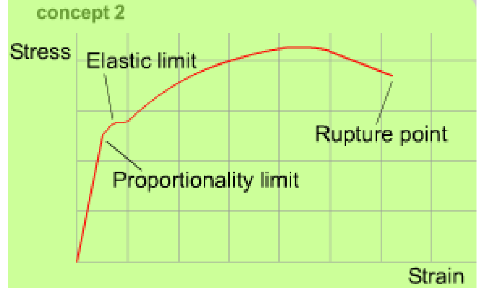
concept 1

Stress and strain
Stress: force per unit area
Strain: fractional change in dimension due to stress
Modulus of elasticity: relates stress, strain for given material

Stress: External force applied per unit area.

Strain: Fractional change in dimension due to stress.

concept 2



Stress-strain graphs
Show behavior of material
· proportionality limit
· elastic limit
· rupture point

Above, you see a machine that tests the behavior of materials under stress. Stress is the external force applied **per unit area** that causes deformation of an object. The machine above stretches the rod, increasing its length. Although we introduce the topic of stress and strain by discussing changes in length, stress can alter the dimensions of objects in other fashions as well.

Strain measures the **fractional** change in an object's dimensions: the object's change in dimension divided by its original dimension. This means that strain is dimensionless. For instance, if a force stretches a two-meter rod by 0.001 meters, the resulting strain is 0.001 meters divided by two meters, which is 0.0005.

Stress measures the force applied per unit area. If a rod is being stretched, the area equals the surface area of the end of the rod, called its cross-sectional area. If the force is being applied over the entirety of an object, such as a ball submerged in water, then the area is the entire surface area of the ball.

For a range of stresses starting at zero, the strain of a material (fractional change in size) is linearly proportional to the stress on it. When the stress ends, the material will resume its original shape: That is, it is elastic. The *modulus of elasticity* is a proportionality constant, the ratio of stress to strain. It differs by material, and depends upon what kind of stress is applied.

Values for modulus of elasticity apply only for a range of stress values. Here, we specifically use changes in length to illustrate this point. If you stress an object beyond its material's *proportionality limit*, the strain ceases to be linearly proportional to stress. A "mild steel" rod will exhibit a linear relation between stress and strain for a stress on it up to about 230,000,000 N/m^2. This is its proportionality limit.

Beyond a further point called the *elastic limit*, the object becomes permanently deformed and will not return to its original shape when the stress ceases to be applied. You have exceeded its *yield strength*. At some point, you pull so far it ruptures. At that *rupture point*, you have exceeded the material's *ultimate strength*.

Concept 2 shows a stress-strain graph typical of a material like soft steel (a type of steel that is easily cut or bent) or copper. Graphs like these are commonly used by engineers. They are created with testing machines similar to the one above by carefully applying force to a material over time. The graph plots the minimum stress required to achieve a certain amount of strain, which in this case is measured as the lengthening of a rod of the material. (We say "minimum stress" because the strain may vary depending on whether a given amount of stress is applied suddenly or is slowly increased to the same value.)

You may notice that the graph has an interesting property: Near the end, the curve flattens and its slope decreases. Less stress is required to generate a certain strain. The material has become *ductile*, or stretchy like taffy, and it is easier to stretch than it was before. Soft steel and copper are ductile. Other materials will reach the rupture point without becoming stretchy; they are called *brittle*. Concrete and glass are two brittle materials, and hardened steel is more brittle than soft steel.

Tensile stress: A stress that stretches.

On the right, you see a rod being stretched. Tensile forces cause materials to lengthen. Tensile stress is the force per cross-sectional unit area of the object. Here, the cross-sectional area equals the surface area of the end of the rod. The strain is measured as the rod's change in length divided by its initial length.

Young's modulus equals tensile stress divided by strain. The letter Y denotes Young's modulus, which is measured in units of newtons per square meter. The value for Young's modulus for various materials is shown in Concept 2. Note the scale used in the table: billions of newtons per square meter.

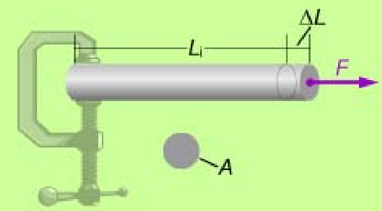
To correctly apply the equation in Equation 1 on the right, you must be careful with the definitions of stress and strain. First, stress is force **per unit area**. Second, strain is the **fractional** change in size. The relevant area for a rod in calculating tensile stress is the area of its end, as the diagrams on the right reflect.

The equation on the right can be used for compression as well as expansion. When a material is compressed, ΔL is the decrease in length. For some materials, Young's modulus is roughly the same for compressing (shortening) as for stretching, so you can use the same modulus when a *compressive force* is applied.

In addition to supplying the values for Young's modulus, we supply values for the yield strength (elastic limit) of a few of the materials. The table can give you a sense of why certain materials are used in certain settings. Steel, for example, has both a high Young's modulus and high yield strength. This means it requires a lot of stress to stretch steel elastically, as well as a lot of stress to deform it permanently.

Some materials have different yield strengths for compression and tension. Bone, for instance, resists compressive forces better than tensile forces. The value listed in the table is for compressive forces.

concept 1



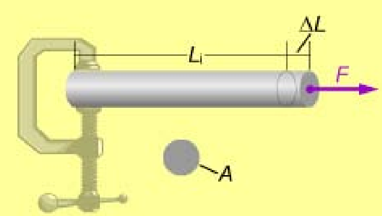
Tensile stress
Causes stretching/compression along a line
Stress: force per cross-section unit area
Strain: fractional change in length

concept 2

	Young's modulus 10^9 (N/m^2)	Yield strength 10^7 (N/m^2)
Aluminum alloys	70	3 to 50
Bone	14	10
Brass	110	6 to 55
Concrete	21	
Copper	108	3 to 40
Glass	70	
Stainless steel	190	20 to 24
Titanium	100	83

Young's modulus
Relates stress, strain

equation 1

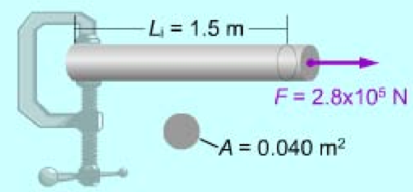


Young's modulus

$$\frac{F}{A} = Y\frac{\Delta L}{L_i}$$

F = force
A = cross-sectional area
Y = Young's modulus (N/m^2)
ΔL = change in length
L_i = initial length

example 1



How much does this aluminum rod stretch under the force?

$$\frac{F}{A} = Y\frac{\Delta L}{L_i}\text{, so } \Delta L = \frac{L_i F}{YA}$$

$$\Delta L = \frac{(1.5\text{ m})\,(2.8\times10^5\text{ N})}{(70\times10^9\text{ N/m}^2)(0.040\text{ m}^2)}$$

$$\Delta L = 1.5\times10^{-4}\text{ m}$$

12.13 - Volume stress

Volume stress: A stress that acts on the entire surface of an object, changing its volume.

The effect of volume stress on a Styrofoam cup submerged 7875 feet underwater (with an un-stressed cup also shown for comparison). The stress on the cup exceeded its elastic limit: It is permanently deformed.

Tensile stress results in a change along a single dimension of an object. A volume stress is one that exposes the entire surface of an object to a force. The force is assumed to be perpendicular to the surface and uniform at all points.

One way to exert volume stress on an object is to submerge it in a fluid (a liquid or a gas). For example, submersible craft that visit the wreck of the Titanic travel four kilometers below the surface, experiencing a huge amount of volume stress on their hulls.

In all cases, stress is force per unit area, which is also the definition of pressure. With volume stresses, the term pressure is used explicitly, since the pressure of fluids is a commonly measured property.

Volume strain is measured as a fractional change in the volume of an object. The modulus of elasticity that relates volume stress and strain is called the *bulk modulus*, and is represented with the letter B.

The equation in Equation 1 states that the change in pressure equals the bulk modulus times the strain. The negative sign means that an **increase** in pressure results in a **decrease** in volume. Unlike the equation for tensile stress, this equation does not have an explicit term for area, because the pressure term already takes this factor into account. Notice that the equation is stated in terms of the **change** in pressure.

At the right is a table that lists values for the bulk modulus for some materials. To give you a sense of the deformation, the increase in pressure at 100 meters depth of water, as compared to the surface, is about 1.0×10^6 N/m^2. The volume of water will be reduced by 0.043% at this depth; steel, only 0.00063%.

concept 1

Volume stress

Pressure of fluids alters volume
Stress: pressure (force per unit area)
Strain: fractional change in volume

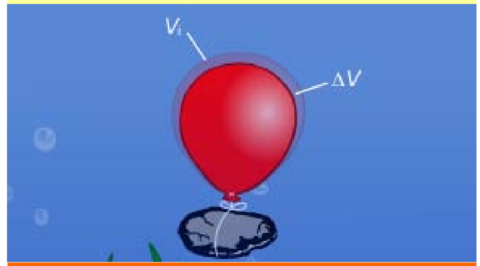
At a depth of 11 km, approximately the maximum depth of the Earth's oceans, the increased pressure is 1.1×10^8 N/m^2. At this depth, a steel ball with a radius of 1.0 meter will compress to a radius of 0.997 m.

concept 2

	Bulk modulus (10^9 N/m^2)
Aluminum alloys	69
Air	0.000144
Brass	110
Copper	120
Glass	37
Mercury	24.8
Saltwater	2.34
Stainless steel	160

Bulk modulus
Relates stress, strain

equation 1



Bulk modulus

$$\Delta P = -B\frac{\Delta V}{V_i}$$

P = pressure (force per unit area)
B = bulk modulus (N/m^2)
V = volume
Pressure units: pascals (Pa = N/m^2)

example 1

The increase in water pressure is 3.2×10^3 Pa. The balloon's initial volume was 0.50 m^3. What is it now?

$$\Delta P = -B\frac{\Delta V}{V_i}, \text{ so } \Delta V = -\frac{V_i \Delta P}{B}$$

$$\Delta V = -\frac{(0.50\ \text{m}^3)(3.2\times10^3\ \text{Pa})}{144{,}000\ \text{N/m}^2}$$

$\Delta V = -0.011$ m^3
$V_f = V_i + \Delta V = 0.50 - 0.011$ m^3
$V_f = 0.49$ m^3

Shear stress: A stress that creates a "slant."

On the right, you see a shearing force. A hand pushes on the block of jello, making it "slant" to the right. This force is parallel to the top surface. An equal and opposing force of friction provided by the plate underneath prevents the block from sliding. These two forces cause shearing. When viewed from the side, the rectangular shape of the block is deformed into a parallelogram by the shearing.

To compute how much an object deforms under a shear stress, we use the equation shown in Equation 1 on the right. As usual, it states that stress equals a modulus times strain.

Stress equals force divided by surface area. With a shearing force, the area is computed from the surface **parallel** to the force being applied. In the example to the right, that is the top surface of the block of gelatin. Why is it this area that is used? Imagine you are holding something in place with a bolt. If you were afraid it would shear off, it would be a better choice to use a wider bolt, instead of a longer one.

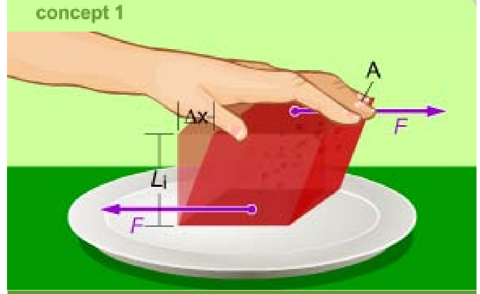
Strain is always a dimensionless ratio. With a shearing stress, strain is a bit more complicated to define than it is for tensile stress. The change depends on how much the top of the block moves sideways, in the direction of the force. The strain associated with shear modulus is the horizontal displacement divided by the height of the block, which changes very little during the shear deformation. This is reflected in the illustration to the right.

S represents the *shear modulus* that relates stress and strain. Like other moduli, the shear modulus has units of newtons per square meter.

On the right, you also see a table that provides the values for the shear modulus for various materials.

This section provides an introduction to shearing stresses. Engineers who must deal with structures that interact dynamically under stress face more complicated conceptual and practical challenges.

concept 1



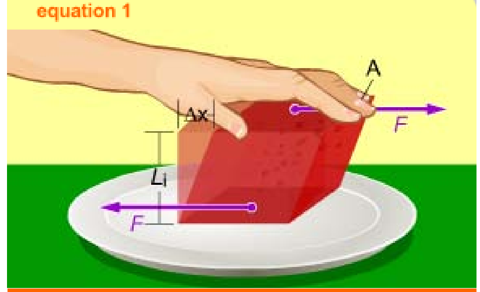
Shear stress
Shearing force: parallel to surface
Stress: force per unit area
Strain: displacement / height

concept 2

	Shear modulus (10^9 N/m^2)
Aluminum alloys	27
Brass	41
Copper	40
Glass	31
Stainless steel	73
Titanium	45

Shear modulus
Relates stress, strain

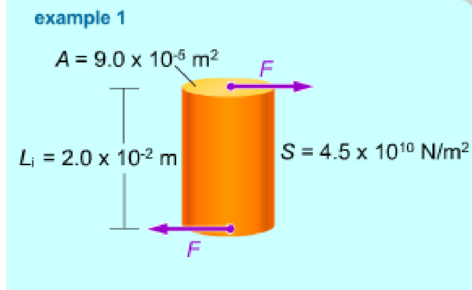
equation 1



Shear modulus

$$\frac{F}{A} = S\frac{\Delta x}{L_i}$$

F = force
A = area of surface parallel to force
S = shear modulus (N/m^2)
Δx = displacement of moved surface
L_i = height of object



What force is required to move the top surface of the titanium cylinder 1.1×10^{-3} m horizontally?

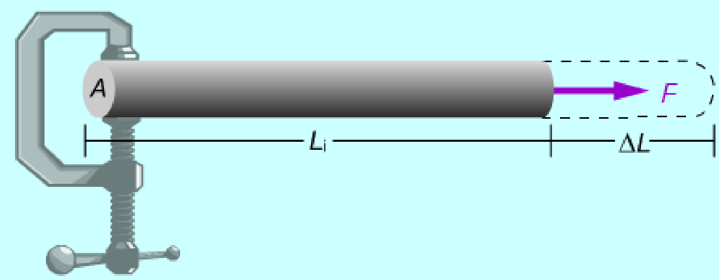
$$\frac{F}{A} = S\frac{\Delta x}{L_i}\text{, so } F = SA\frac{\Delta x}{L_i}$$

$$F =$$

$$(4.5\times10^{10}\,\text{N/m}^2)(9.0\times10^{-5}\,\text{m}^2)\frac{1.1\times10^{-3}\,\text{m}}{2.0\times10^{-2}\,\text{m}}$$

$$F = 2.2\times10^{5}\,\text{N}$$

12.15 - Interactive checkpoint: stress and strain



A 3.50 meter-long rod, composed of a titanium alloy, has a cross-sectional area of 4.00×10^{-4} m^2. It increases in length by 0.0164 m under a force of 2.00×10^{5} N. What is the stress on the rod? What is the strain? What is Young's modulus for this titanium alloy?

Answer:

$stress =$ ______ N/m^2

$strain =$ ______

$Y =$ ______ N/m^2

12.16 - Gotchas

Stress equals force. No, stress is always a measure of force per unit surface area.

A force stretches a rod by 0.01 meters, so the strain is 0.01 meters. No, strain is always the fractional change (a dimensionless ratio). To calculate the strain, you need to divide this change in length by the initial length of the rod. This is not stated here, so you cannot determine the strain without more information.

12.17 - Summary

An object is in equilibrium when there are no net forces or torques on it. Static equilibrium is a special case where there is also no motion.

The center of gravity is the average location of the weight of an object. The force of gravity effectively acts on the object at this point. The concept of center of gravity is very similar to the concept of center of mass. The two locations differ only when an object is so large that the pull of gravity varies across it.

Elasticity refers to an object's shape changing when forces are applied to it. Objects are called elastic when they return to their original shape as forces are removed.

Related to elasticity are stress and strain. Stress is the force applied to an object per unit area. The area used to determine stress depends on how the force is applied. Strain is the fractional change in an object's dimensions due to stress.

The relationship between stress and strain is determined by a material's modulus of elasticity up to the proportionality limit. An object will become permanently deformed if it is stressed past its elastic limit. It will finally break at its rupture point. Ductile materials are easily deformed, while brittle materials tend to break rather than stretch.

Tensile stress is the application of stress causing stretching or compression along a line. Strain under tensile stress is measured as a fractional change in length and stress is measured as the force per unit cross-sectional area. Young's modulus is the tensile stress on a material divided by its strain.

Equations

Static equilibrium

$$\Sigma F_{\mathrm{x}} = 0,\ \Sigma F_{\mathrm{y}} = 0,\ \Sigma\tau = 0$$

Tensile stress

$$\frac{F}{A} = Y\frac{\Delta L}{L_{\mathrm{i}}}$$

Volume stress

$$\Delta P = -B\frac{\Delta V}{V_{\mathrm{i}}}$$

Shear stress

$$\frac{F}{A} = S\frac{\Delta x}{L_{\mathrm{i}}}$$

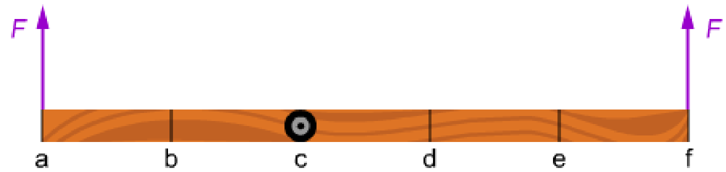
Volume stress acts over the entire surface of an object. The amount of volume stress is calculated as the force per unit surface area, while the strain is the fractional change in volume. The bulk modulus relates volume stress and strain.

Shear stress is a stress parallel to a surface. It causes a strain that is seen as a slanting deformation. The stress is measured as the force per unit area parallel to the force. The strain is computed as the displacement divided by the height of an object perpendicular to the face experiencing the force. The shear modulus relates shear stress and strain for a particular material.

Chapter 12 Problems

Conceptual Problems

C.1 The sketch shows a rod divided into five equal parts. The rod has negligible mass and a fixed pivot at point c. An upward force of magnitude *F* is applied at point a, and an identical force is applied at point f, as shown. At what point on the rod could you apply a third force of the same magnitude *F*, perpendicular to the rod, either upward or downward, so that the net torque on the rod is zero? Check all of the points for which this is possible.

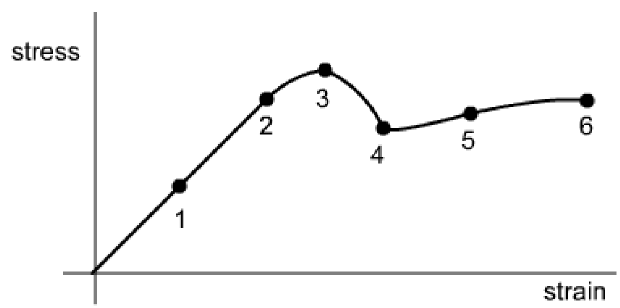


- ☐ a
- ☐ b
- ☐ c
- ☐ d
- ☐ e
- ☐ f

C.2 Use principles of equilibrium to describe how the gymnast is able to hold his body in this position.

C.3 The Space Needle in Seattle, Washington is over 500 feet tall, but the center of gravity of the complete structure is only 5 feet off the ground. This is achieved with a huge concrete foundation that weighs more than the Space Needle itself. What is the advantage of having a center of gravity so low?

C.4 (a) Which point is the proportionality limit? (b) Which point is the rupture point? (c) Which point is the elastic limit?



(a) i. Point 1
ii. Point 2
iii. Point 3
iv. Point 4
v. Point 5
vi. Point 6

(b) i. Point 1
ii. Point 2
iii. Point 3
iv. Point 4
v. Point 5
vi. Point 6

(c) i. Point 1
ii. Point 2
iii. Point 3
iv. Point 4
v. Point 5
vi. Point 6

C.5 Suppose you have four rods made of four different materials-aluminum, copper, steel and titanium. They all have the same dimensions, and the same force is applied to each. Order the rods from smallest to largest by how much each stretches.

i. Aluminum ,	i. Aluminum ,	i. Aluminum ,	i. Aluminum
ii. Copper	ii. Copper	ii. Copper	ii. Copper
iii. Steel	iii. Steel	iii. Steel	iii. Steel
iv. Titanium	iv. Titanium	iv. Titanium	iv. Titanium

Section Problems

Section 1 - Static equilibrium

1.1 Two children sit 2.60 m apart on a very low-mass horizontal seesaw with a movable fulcrum. The child on the left has a mass of 29.0 kg, and the child on the right has a mass of 38.0 kg. At what distance, as measured from the child on the left, must the fulcrum be placed in order for them to balance?

__________ m

1.2 A boy carries a sack on one end of a very light stick that is balanced on his shoulder, at an angle of 22.0 degrees up from the horizontal. The mass of the sack is 7.00 kg, and it sits 1.20 m from his shoulder. If his hand is 0.350 m from his shoulder, on the other end of the stick, what is the magnitude of the downward force the boy exerts with his hand?

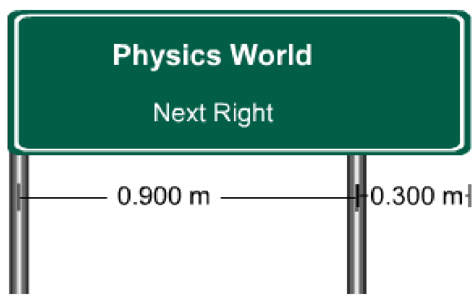
__________ N

1.3 A 3.6-meter long horizontal plank is held up by two supports. One support is at the left end, and the other is 0.80 m from the right end. The plank has uniform density and has a mass of 40 kg. How close can a 70 kg person stand to the unsupported end before causing the plank to rotate?

__________ m

1.4 The mass of the sign shown is 28.5 kg. Find the weight supported by (a) the left support and (b) the right support.

(a) __________ N
(b) __________ N



1.5 A horizontal log just barely spans a river, and its ends rest on opposite banks. The log is uniform and weighs 2400 N. If a person who weighs 840 N stands one fourth of the way across the log from the left end, how much weight does the bank under the right end of the log support?

__________ N

1.6 A person who weighs 620 N stands at *x* = 5.00 m, right on the end of a long horizontal diving board that weighs 350 N. The diving board is held up by two supports, one at its left end at *x* = 0, and one at the point *x* = 2.00 m. (a) What is the force exerted on the support at *x* = 0? (b) What is the force acting on the other support? (Use positive to indicate an upward force, negative for a downward force.)

(a) __________ N
(b) __________ N

1.7 Two brothers, Jimmy and Robbie, sit 3.00 m apart on a horizontal seesaw with its fulcrum exactly midway between them. Jimmy sits on the left side, and his mass is 42.5 kg. Robbie's mass is 36.5 kg. Their sister Betty sits at the exact point on the seesaw so that the entire system is balanced. If Betty is 29.8 kg, at what location should she sit? Take the fulcrum to be the origin, and right to be positive. Assume that the mass of the seesaw is negligible.

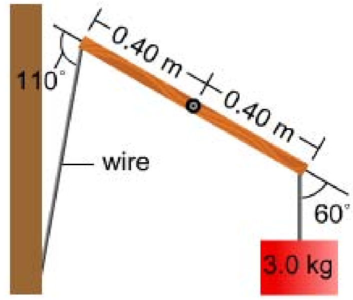
__________ m

1.8 If a cargo plane is improperly loaded, the plane can tip up onto its tail while it rests on the runway (this has actually happened). Suppose a plane is 45.0 m long, and weighs 1.20×10^6 N. The center of mass of the plane is located 21.0 m from the nose. The nose wheel located 3.50 m from the nose and the main wheels are 25.0 m from the nose. Cargo is loaded into the back end of the plane. If the center of mass of the cargo is located 40.0 m from the nose of the plane, what is the maximum weight of the cargo that can be put in the plane without tipping it over (that is, so that the plane remains horizontal)?

__________ N

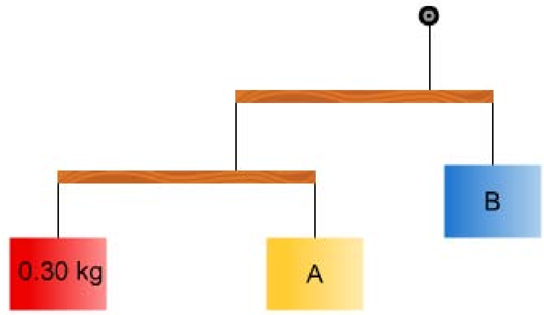
1.9 The picture shows a mass hanging from a rod that is free to pivot about the point shown. The other end of the rod is anchored to the wall with a wire. The system is in equilibrium. What is the tension in the wire?

__________ N



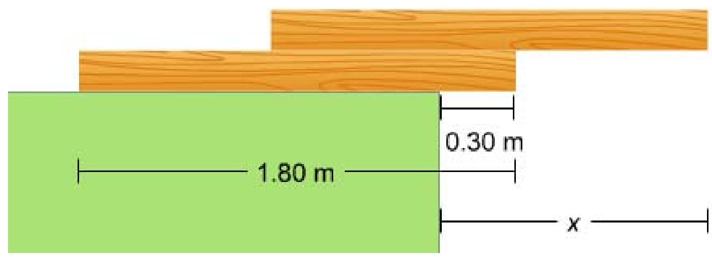
1.10 The sketch shows a mobile in equilibrium. Each of the rods is 0.16 m long, and each hangs from a supporting string that is attached one fourth of the way across it. The mass of each rod is 0.10 kg. The mass of the strings connecting the blocks to the rods is negligible. What is the mass of (a) block A? (b) block B?

(a) __________ kg
(b) __________ kg



1.11 Two identical 1.80 m long boards just barely balance on the edge of a table, as shown in the figure. What is the distance *x*?

__________ m



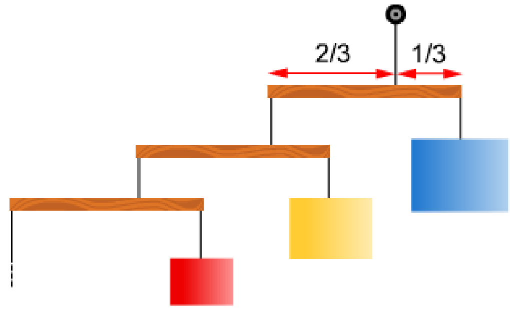
1.12 The wheelbarrow shown has a mass of 35.0 kg without the wheel. When it is empty, the center of mass is 0.400 m from the axle. A load with a mass of 55.0 kg is put in the wheelbarrow, 0.600 m from the axle. Olivia holds the handles of the wheelbarrow, 1.20 m from the axle. She lifts the handles so that they make a 25.0° angle with the ground. What is the upward force that Olivia applies to the handles in order to hold the wheelbarrow in that position?

__________ N

1.13 A string of length 1.41 m is strung loosely between two points that are separated horizontally by 1.24 m and vertically by 0.515 m (these points are 1.34 m apart). A 1.54 kg mass is attached to the string 1.14 m from the lower end, as measured along the string. This bends the string into two "straight" line segments at different angles. What are the tensions in (a) the lower part of the string and (b) the upper part of the string?

(a) __________ N
(b) __________ N

1.14 (a) A mobile is constructed by connecting a string to a massless rod one-third of the way from the end. A block is attached to each end of the rod. The mobile is in static equilibrium. What is the mass of the block hanging from the end of the rod farther from the string, as a fraction of the mobile's total mass *m*? (b) Suppose that the block hanging from the far end is replaced by another massless rod with two dependent blocks, having the same net mass as the single block they replace. The string supporting the lower rod is attached one-third of the way from its end. What is the mass of the block on the end of the lower rod that is farther from the supporting string, as a fraction of the mobile's total mass *m*? (c) Now suppose the mobile consists of a chain of *n* rods connected in the same manner. What is the mass of the block on the far end of the lowest rod as a fraction of the total mass? Express your answer in terms of *n*.



(a) ○ 1/9 ○ 2/9 ○ 1/3 ○ 1/2 ○ 2/3
(b) ○ 1/27 ○ 1/9 ○ 4/27 ○ 2/9 ○ 1/4
(c) ○ $1/3^n$ ○ $1/2^n$ ○ $2/3^n$ ○ $2/3^n$ ○ $1/n^3$

Section 3 - Center of gravity

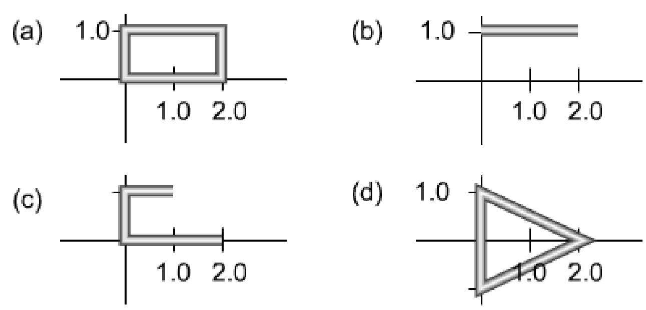
3.1 A weightlifter has been given a barbell to lift. One end has a mass of 5.5 kg while the other end has a mass of 4.7 kg. The bar is 0.20 m long. (Consider the bar to be massless, and assume that the masses are thin disks, so that their centers of mass are at the ends of the bar.) How far from the heavier end should she hold the bar so that the weight feels balanced?

__________ m

3.2 A 0.65 m rod with uniform mass distribution runs along the *x* axis with its left end at the origin. A 1.8 m rod with uniform mass distribution runs along the *y* axis with its top end at the origin. Find the coordinates of the center of gravity for this system.

(__________ , __________) m

3.3 A length of uniform wire is cut and bent into the shapes shown. Find the location of the center of gravity of each shape. In each instance, consider the corners of the shape to be located at integer coordinates.



(a) (__________ , __________)
(b) (__________ , __________)
(c) (__________ , __________)
(d) (__________ , __________)

3.4 Three beetles stand on a grid. Two beetles have the same weight, *W*, and the third beetle weighs 2*W*. (a) The lighter beetles are located at (1.00, 0) and (0, 2.00), and the heavier beetle is at (3.00, 1.00). Find the coordinates of the center of gravity of the beetles. (b) If the heavy beetle moves to (1.00, 1.00), what is the new location of the center of gravity?

(a) (__________ , __________)
(b) (__________ , __________)

3.5 A woman with weight 637 N lies on a bed of nails. The bed has a weight of 735 N and a length of 1.72 m. The bed is held up by two supports, one at the head and one at the foot. Underneath each support is a scale. When the woman lies in the bed, the scale at the foot reads 712 N. How far is the center of gravity of the system from the foot of the bed of nails?

__________ m

3.6 (a) An empty delivery truck weighs 5.20×10^5 N. Of this weight, 3.20×10^5 N is on the front wheels. The distance between the front axle and the back axle is 4.10 m. How far is the center of gravity of the truck from the front wheels? (b) Now the delivery truck is loaded with a 2.40×10^5 N shipment, 2.60 m from the front wheels. Now how far is the center of gravity from the front wheels?

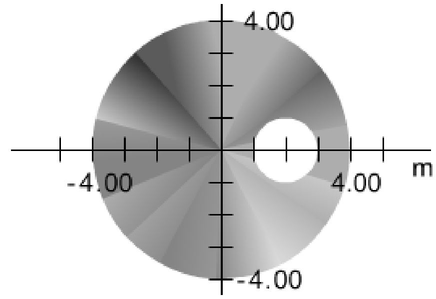
(a) __________ m
(b) __________ m

3.7 A skateboarder stands on a skateboard so that 62% of her weight is located on the front wheels. If the distance between the wheels is 0.85 m, how far is her center of gravity from the front wheels? Ignore the weight of this rather long skateboard.

__________ m

3.8 The figure shows a uniform disk with a circular hole cut out of it. The disk has a radius of 4.00 m. The hole's center is at 2.00 m, and its radius is 1.00 m. What is the *x*-coordinate of the center of gravity of the system?

__________ m



Section 4 - Sample problem: a leaning ladder

4.1 A ladder leans against a wall making a 55.0° angle to the floor. The ladder is 4.50 m long, and weighs 415 N. The wall is frictionless and so is the floor. A horizontal wire is attached to the base of the ladder and attached to the wall. (a) What is the tension in the wire? (b) A person who weighs 655 N stands on a rung of the ladder located 2.00 m from its lower end. What is the new tension in the wire?

(a) __________ N
(b) __________ N

4.2 A ladder is made in the shape of the letter A. Treat the two sides of the ladder as identical uniform rods, each weighing 455 N, with a length of 3.60 m. A frictionless hinge connects the two ends at the top, and a horizontal wire, 1.20 m long, connects them at a distance 1.40 m from the hinge, as measured along the sides. The ladder rests on a frictionless floor. What is the tension in the wire?

__________ N

Section 7 - Interactive problem: achieve equilibrium

7.1 Use the data and the diagram in the interactive problem in this section to answer the following question. If the pivot is placed at the left end of the rod, what is (a) the magnitude and (b) direction of the unknown force? (c) How far should it be placed from the pivot? Use the simulation to test your answer.

(a) __________ N
(b) i. Positive *x*
ii. Negative *x*
iii. Positive *y*
iv. Negative *y*
(c) __________ m

Section 11 - Stress and strain

11.1 (a) Suzy the elephant weighs 65,000 N. The cross-sectional area of each foot is 0.10 m^2. When she stands on all fours, what is the average stress on her feet? (b) Suzanne the stockbroker weighs 610 N. She wears a pair of high-heeled shoes whose heels each have a cross-sectional area of 1.5×10^{-4} m^2 in contact with the floor. If she stands with all her weight on her heels, what is the stress on the heels of the shoes?

(a) __________ N/m^2
(b) __________ N/m^2

Section 12 - Tensile stress

12.1 A 85.0 kg window washer hangs down the side of a building from a rope with a cross-sectional area of 4.00×10^{-4} m^2. If the rope stretches 0.740 cm when it is let out 7.50 m, how much will it stretch when it is let out 26.0 m?

__________ m

12.2 A building with a weight of 4.10e+7 N is built on a concrete foundation with an area of 1300 m^2 and an (uncompressed) height of 2.82 m. By what vertical distance does the foundation compress?

__________ m

12.3 Many gyms have an exercise machine called a vertical leg press. To use this machine, you lie on your back and press a weight upward with your feet until your legs are perpendicular to the ground. Find the amount the femur bone compresses when a person with a 0.385 m femur lifts 525 N with a vertical leg press, using one leg, until the leg is fully extended straight up. Assume that the cross sectional area of the femur is 5.30×10^{-4} m^2.

__________ m

12.4 (a) An engineer is designing a bridge with four supports made of cylinders of concrete. The cylinders are 3.5 m tall, and they cannot compress by more than 1.6×10^{-4} m. If the bridge must be able to hold 7.5×10^{6} N, what should be the radius of the cylinders? (b) Suppose the restriction on the amount of compression is removed. If the yield strength of the concrete is 5.7×10^{6} N/m^2, what is the smallest radius that will prevent the concrete from deforming under the weight?

(a) __________ m
(b) __________ m

12.5 A climber with a mass of 75 kg is attached to a 6.5 m rope with a cross sectional area of 3.5×10^{-4} m^2. When the climber hangs from the rope, it stretches 3.4×10^{-3} m. What is the Young's modulus of the rope?

__________ N/m^2

12.6 A helicopter is working to bring water to a forest fire. A steel cable hangs from the helicopter with a large bucket on the end. The cable is 8.5 m long and has a cross-sectional area of 4.0×10^{-4} m^2. Together, the bucket and the water inside it weigh 3.0×10^{4} N. (a) How much will the cable stretch when the helicopter is hovering? (b) If the yield strength of the steel in the cable is 2.1×10^{8} N/m^2, at what rate can the helicopter accelerate upward before the cable deforms permanently?

(a) __________ m
(b) __________ m/s^2

12.7 Two rods with identical dimensions are placed end to end to form a new rod. If one of the smaller rods is aluminum alloy and the other is titanium, what is the effective Young's modulus of the large rod?

__________ N/m^2

Section 13 - Volume stress

13.1 The deepest point in the seven seas is the Marianas Trench in the Pacific Ocean. The pressure in the deepest parts of the Marianas Trench is
1.1×10^{8} Pa. Pressure at the surface of the ocean is 1.0×10^{5} Pa. If a mass of salt water has a volume of 1.6 m^3 at the surface of the ocean, what will be its volume at the bottom of the Marianas Trench?

__________ m^3

13.2 If a spherical glass marble has a radius of 0.00656 m at 1.02×10^{5} Pa, at what pressure will it have a radius of 0.00650 m?

__________ Pa

13.3 The pressure on the surface of Venus is about 9.0×10^{6} Pa, and the pressure on the surface of Earth is 1.0×10^{5} Pa. What would be the volume strain on a solid stainless steel sphere if it were moved from Earth to Venus?

____

13.4 A sealed, expandable plastic bag is filled with equal volumes of saltwater and mercury. The volume of the bag is 0.120 m^3 at a pressure of 1.01×10^{5} N/m^2. What is the change in volume of the bag when the pressure increases to 1.09×10^{5} N/m^2? Assume that the bag exerts no force.

__________ m^3

13.5 A solid aluminum-alloy sphere is in a vacuum chamber where the pressure is 1.3×10^{-7} Pa. It is moved from the chamber into a room with air pressure 1.0×10^{5} Pa. By what fraction, $\Delta r/r_i$, does the radius of the sphere change when it is moved out of the vacuum? (Be sure to indicate sign.)

____

13.6 The volume strain equation is not based on physical laws. It is an empirically derived approximation that has been experimentally verified for a wide range of materials. It is accurate when the change in pressure ΔP is numerically small compared to the bulk modulus B. (Physicists say that ΔP must be "dominated" by B, a condition which is written $\Delta P << B$). Since bulk moduli are numerically very large, this is a condition that is ordinarily easily satisfied.
Prove the necessity of $\Delta P << B$ in the following steps. (a) Suppose an object of volume V_1 is subjected to a change of pressure ΔP_{12} that compresses (or expands) the object to a new volume V_2. Show that $V_2 = V_1(1 - \Delta P_{12}/B)$. (b) Then suppose the object is subjected to a further change of pressure ΔP_{23} that compresses (or expands) the object to a new volume V_3. Derive a similar formula expressing V_3 in terms of V_2. (c) Combine parts a and b to express V_3 in terms of V_1 when the pressure changes are equal and opposite, that is, when $\Delta P_{23} = -\Delta P_{12}$. (d) Since the changes described in part c return the object to its original condition, then V_3 should be the same as V_1. Why does this require that $\Delta P_{12} << B$?
The domination requirement prevents the volume strain equation from being applied to air if the pressure difference ΔP is on the order of atmospheric pressure. The bulk modulus of air is only $B = 1.44\times10^{5}$ N/m^2, while one atmosphere of pressure is $P = 1.013\times10^{5}$ Pa, which is numerically of the same order of magnitude. In a later chapter the *ideal gas law* will provide you with a reliable and exact way to relate the volume of a gas to the pressure on it, over a wide range of values.

Section 14 - Shear stress

14.1 A 110 kg sculpture is attached to the wall by 4 identical bolts inserted into horizontally-drilled holes in the wall. If each bolt has a circular cross-sectional area of 5.50×10^{-5} m^2, what is the shear stress on the bolts?

__________ N/m2

14.2 A brass brick, 0.35 m long, 0.29 m wide and 0.25 m tall, is subjected to two equal but opposite shear forces, one on the top and one on the bottom of the brick. The shear forces cause the brick to deform so that the angle between the side of the brick and the vertical is 0.065°. What is the magnitude of each of the shear forces?

__________ N

14.3 A glass cube with sides of 0.61 m is subjected to a shear force of 1.5×10^6 N. (a) What is the shear stress on the cube? (b) What is the strain? (c) By how much is the top of the cube displaced?

(a) __________ N/m2
(b) __________
(c) __________ m

14.4 A diving board is horizontal when there is no one on it. When a swimmer with a mass of 55 kg stands on the end, the tip of the board displaces by 0.15 m. If a 95 kg swimmer stands on the end of the board, by how much will the tip displace?

__________ m

Additional Problems

A.1 A kung-fu master holds a 3.00 kg, 3.60 m long, uniform bamboo staff at a 30.0° angle down toward the ground in front of her. Her left hand is supporting the pole 0.700 m down from the center of the pole, and acting as a pivot point. Her right hand is 0.800 m down from her left, pressing downward on the lower end of the pole. Find the magnitudes of (a) the upward force exerted by her left hand and (b) the downward force exerted by her right hand in order to keep the pole in static equilibrium.

(a) __________ N
(b) __________ N

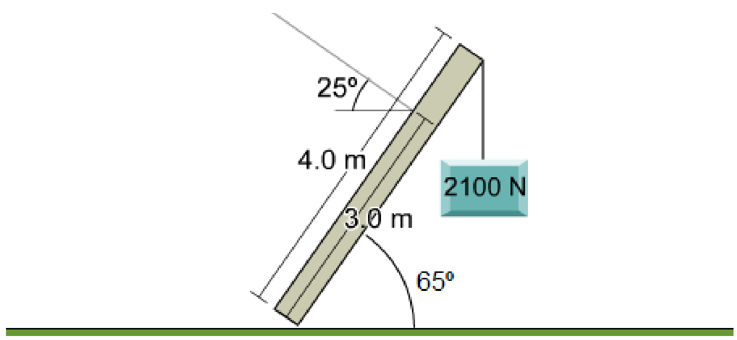
A.2 A squat aluminum-alloy cylinder, weighing 1.2×10^4 N, is 0.30 m tall and has a radius of 0.70 m. The coefficient of static friction between the cylinder and the surface it sits on is 0.80. (a) What is the maximum shear force that can be applied to the top of the cylinder without it moving? (b) What is the horizontal displacement of the top of the cylinder when this force is applied?

(a) __________ N
(b) __________ m

A.3 A slab weighing 15,000 N rests on a frictionless plane. The slab is attached to the plane with four screws, each with a cross-sectional area of 5.0×10^{-3} m^2. What is the shear stress on each screw if the plane is inclined to 32 degrees off the horizontal?

__________ N/m2

A.4 A uniform, 4.0-meter long beam weighing 1200 N is supported by a cable, as shown in the figure. The beam pivots at the bottom, and a 2100 N dead weight hangs from its top. Since the beam is not accelerating, you know that the net torque on it is zero. Find the tension in the cable, which is oriented at 25° from the horizontal i.e., perpendicular to the beam.



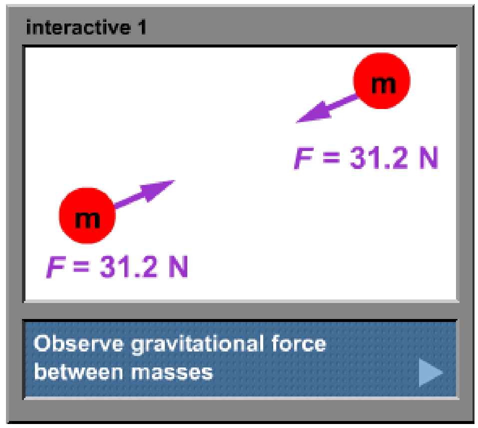
__________ N

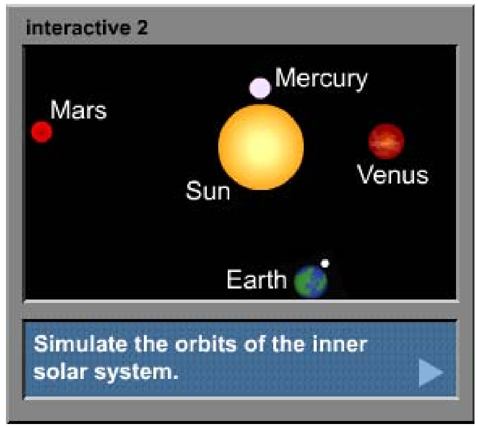
13.0 - Introduction

The topic of gravity has had a starring role in some of the most famous tales in the history of physics. Galileo Galilei was studying the acceleration due to the Earth's gravity when he dropped two balls from the Leaning Tower of Pisa. A theory to explain the force of gravity came to Isaac Newton shortly after an apple fell from a tree and knocked him in the head. Although historians doubt whether these events actually occurred, the stories have come to symbolize how a simple experiment or a sudden moment of insight can lead to important and lasting scientific progress.

Galileo is often said to have dropped two balls of different masses from the tower so that he could see if they would land at the same time. Most scientists of his era would have predicted that the heavier ball would hit the ground first. Instead, the balls landed at the same instant, showing that the acceleration due to gravity is a constant for differing masses. While it is doubtful that Galileo actually dropped balls off that particular tower, his writings show that he performed many experiments studying the acceleration due to gravity.

In another famous story, Newton formed his theory of gravity after an apple fell and hit him on the head. At least one person (the daughter of French philosopher Voltaire) said that Newton mentioned that watching a falling apple helped him to comprehend gravity. Falling apple or no, his theory was not the result of a momentary insight; Newton pondered the topic of gravity for decades, relying on the observations of contemporary astronomers to inform his thinking. Still, the image of a scientist deriving a powerful scientific theory from a simple physical event has intrigued people for generations.





The interactive simulations on the right will help you begin your study of gravity.

Interactive 1 permits you to experiment with the gravitational forces between objects. In its control panel, you will see five point masses. There are three identical red masses of mass *m*. The green mass is twice as massive as the red, and the blue mass is four times as massive.

You can start your experimentation by dragging two masses onto the screen. The purple vectors represent the gravitational forces between them. You can move a mass around the screen and see how the gravitational forces change. What is the relationship between the magnitude of the forces and the distances between the masses?

Experiment with different masses. Drag out a red mass and a green mass. Do you expect the force between these two masses to be smaller or larger than it would be between two red masses situated the same distance apart? Make a prediction and test it. You can also drag three or more masses onto the screen to see the gravitational force vectors between multiple bodies. There is yet more to do: You can also press GO and see how the gravitational forces cause the masses to move.

The other major topic of this chapter is orbital motion. The force of gravity keeps bodies in orbit. Interactive 2 is a reproduction of the inner part of our solar system. It shows the Sun, fixed at the center of the screen, along with the four planets closest to it: Mercury, Venus, Earth and Mars.

Press GO to watch the planets orbit about the Sun. You can experiment with our solar system by changing the position of the planets prior to pressing GO, or by altering the Sun's mass as the planets orbit.

In the initial configuration of this system, the period of each planet's orbit (the time it takes to complete one revolution around the Sun) is proportional to its actual orbital period. Throughout this chapter, as in this simulation, we will often not draw diagrams to scale, and will speed up time. If we drew diagrams to scale, many of the bodies would be so small you could not see them, and taking 365 days to show the Earth completing one revolution would be asking a bit much of you.

13.1 - Newton's law of gravitation

Newton's law of gravitation: The attractive force of gravity between two particles is proportional to the product of their masses and inversely proportional to the square of the distance between them.

Newton's law of gravitation states that there is a force between every pair of particles, of any mass, in the universe. This force is called the gravitational force, and it causes objects to attract one another. The force does not require direct contact. The Earth attracts the Sun, and the Sun attracts the Earth, yet 1.5×10^{11} meters of distance separate the two.

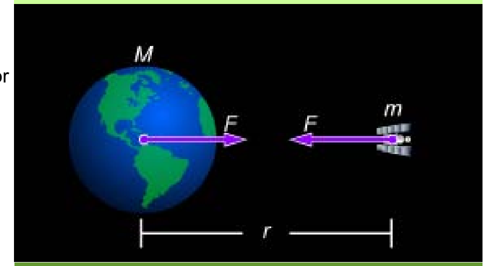
The strength of the gravitational force increases with the masses of the objects, and weakens proportionally to the square of the distance between them. Two masses exert equal but opposite attractive forces on each other. The forces act on a line between the two objects. The magnitude of the force is calculated using the equation on the right. The symbol G in the equation is the *gravitational constant*.

It took Newton about 20 years and some false starts before he arrived at the relationship between force, distance and mass. Later scientists established the value for G, which equals 6.674 2×10^{-11} N·m²/kg². This small value means that a large amount of mass is required to exert a significant gravitational force.

The example problems on the right provide some sense of the magnitude of the gravitational force. First, we calculate the gravitational force the Earth exerts on the Moon (and the Moon exerts on the Earth). Although separated by a vast distance (on average, their midpoints are separated by about 384,000,000 meters), the Earth and the Moon are massive enough that the force between them is enormous: 1.98×10^{20} N.

In the second example problem, we calculate the gravitational force between an 1100-kg car and a 2200-kg truck parked 15 meters apart. The force is 0.00000072 N. When you press a button on a telephone, you press with a force of about one newton, 1,400,000 times greater than this force.

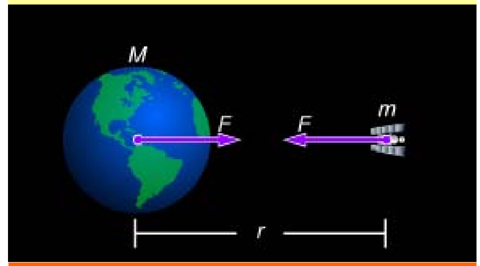
concept 1



Newton's law of gravitation

Gravitational force
· Proportional to masses of bodies
· Inversely proportional to square of distance

equation 1

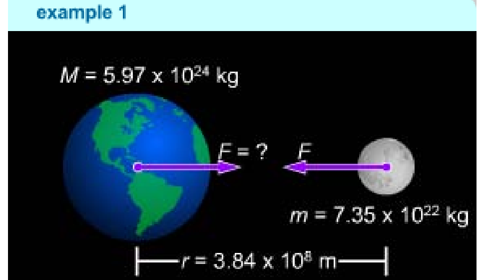


Newton's law of gravitation

$$F = \frac{GMm}{r^2}$$

F = force of gravity
M, m = masses of objects
r = distance between objects
G = gravitational constant
$G = 6.67\times10^{-11}$ N·m²/kg²

example 1



What is the gravitational force between the Earth and the Moon?

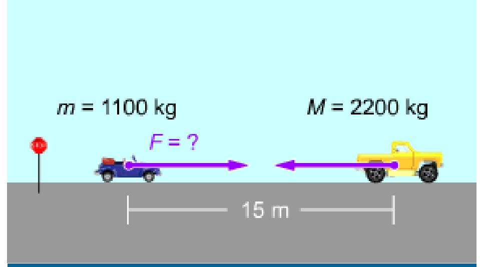
$$F = \frac{GMm}{r^2}$$

$$F = \frac{(6.67\times10^{-11})(5.97\times10^{24}\,\text{kg})(7.35\times10^{22}\,\text{kg})}{(3.84\times10^{8}\,\text{m})^2}$$

$F = 1.98\times10^{20}$ N

Each body attracts the other

example 2



How much gravitational force do the car and the truck exert on each other?

$$F = \frac{GMm}{r^2}$$

$$F = \frac{(6.67\times10^{-11})(1100\text{ kg})(2200\text{ kg})}{(15\text{ m})^2}$$

$$F = 7.2\times10^{-7}\text{ N}$$

13.2 - G and g

Newton's law of gravity includes the gravitational constant G.

In this section, we discuss the relationship between G and g, the rate of freefall acceleration in a vacuum near the Earth's surface.

The value of g used in this textbook is 9.80 m/s^2, an average value that varies slightly by location on the Earth. Both a 10-kg object and a 100-kg object will accelerate toward the ground at 9.80 m/s^2. The rate of freefall acceleration does not vary with mass.

Newton's law of gravitation, however, states that the Earth exerts a stronger gravitational force on the more massive object. If the force on the more massive object is greater, why does gravity cause both objects to accelerate at the same rate?

The answer becomes clear when Newton's second law of motion, $F = ma$, is applied. The acceleration of an object is proportional to the force acting on it, divided by the object's mass. The Earth exerts ten times the force on the 100-kg object that it does on the 10-kg object. But that tenfold greater force is acting on an object with a mass ten times greater, meaning the object has ten times more resistance to acceleration. The result is that the mass term cancels out and both objects accelerate toward the center of the Earth at the same rate, g.

If the gravitational constant G and the Earth's mass and radius are known, then the acceleration g of an object at the Earth's surface can be calculated. We show this calculation in the following steps. The distance r used below is the average distance from the surface to the center of the Earth, that is, the Earth's average radius. We treat the Earth as a particle, acting as though all of its mass is at its center.

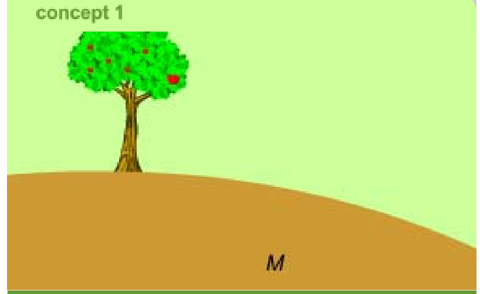
Variables

gravitational constant	$G = 6.67\times10^{-11}\text{ N·m}^2\text{/kg}^2$
mass of the Earth	$M = 5.97\times10^{24}\text{ kg}$
mass of object	m
Earth-object distance	$r = 6.38\times10^6\text{ m}$
acceleration of object	g

Strategy

1. Set the expressions for force from Newton's second law and his law of gravitation equal to each other.
2. Solve for the acceleration of the object, and evaluate it using known values for the other quantities in the equation.

concept 1

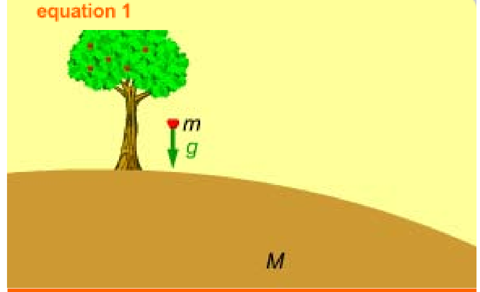


G* and *g

G = gravitational constant everywhere in universe

g = freefall acceleration at Earth's surface

equation 1



G* and *g

$$g = \frac{GM}{r^2}$$

g = free fall acceleration at Earth's surface

G = gravitational constant

M = mass of Earth

r = distance to center of Earth

Physics principles and equations

We will use Newton's second law of motion and his law of gravitation. In this case, the acceleration is g.

$$F = ma = mg, \quad F = \frac{GMm}{r^2}$$

Step-by-step derivation

Here we use two of Newton's laws, his second law ($F = ma$) and his law of gravitation ($F = GMm/r^2$). We use g for the acceleration instead of a, because they are equal. We set the right sides of the two equations equal and solve for g.

Step	Reason
1. $mg = \frac{GMm}{r^2}$	Newton's laws
2. $g = \frac{GM}{r^2}$	simplify
3. $g = \frac{(6.67\times10^{-11}\ \frac{\text{N}\cdot\text{m}^2}{\text{kg}^2})(5.97\times10^{24}\ \text{kg})}{(6.38\times10^{6}\ \text{m})^2}$	substitute known values
4. $g = 9.78\ \text{m/s}^2$	evaluate

Our calculations show that g equals 9.78 m/s^2. The value for g varies by location on the Earth for reasons you will learn about later.

In the steps above, the value for G, the gravitational constant, is used to calculate g, with the mass of the Earth given in the problem. However, if g and G are both known, then the mass of the Earth can be calculated, a calculation performed by the English physicist Henry Cavendish in the late 18th century.

Physicists believe G is the same everywhere in the universe, and that it has not changed since the Big Bang some 13 billion years ago. There is a caveat to this statement: some research indicates the value of G may change when objects are extremely close to each other.

13.3 - Shell theorem

Shell theorem: The force of gravity outside a sphere can be calculated by treating the sphere's mass as if it were concentrated at the center.

Newton's law of gravitation requires that the distance between two particles be known in order to calculate the force of gravity between them. But applying this to large bodies such as planets may seem quite daunting. How can we calculate the force between the Earth and the Moon? Do we have to determine the forces between all the particles that compose the Earth and the Moon in order to find the overall gravitational force between them?

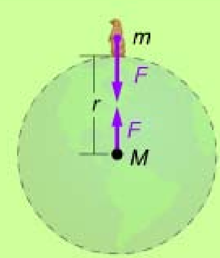
Fortunately, there is an easier way. Newton showed that we can assume the mass of each body is concentrated at its center.

Newton proved mathematically that a uniform sphere attracts an object outside the sphere as though all of its mass were concentrated at a point at the sphere's center. Scientists call this the shell theorem. (The word "shell" refers to thin shells that together make up the sphere and which are used to mathematically prove the theorem.)

Consider the groundhog on the Earth's surface shown to the right. Because the Earth is approximately spherical and the matter that makes up the planet is distributed in a spherically symmetrical fashion, the shell theorem can be applied to it. To use Newton's law of gravitation, three values are required: the masses of two objects and the distance between them. The mass of the groundhog is 5.00 kg, and the Earth's mass is 5.97×10^{24} kg. The distance between the groundhog and the center of the Earth is the Earth's radius, which averages 6.38×10^6 meters.

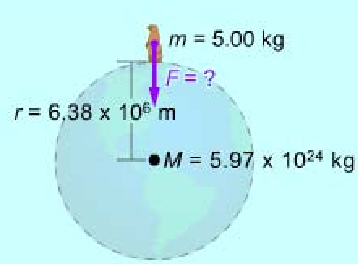
In the example problem to the right, we use Newton's law of gravitation to calculate the gravitational force exerted on the groundhog by the Earth. The force equals the groundhog's weight (mg), as it should.

concept 1



The shell theorem
Consider sphere's mass to be concentrated at center
· r is distance between centers of spheres

example 1



How much gravitational force does the Earth exert on the

groundhog?

$$F = \frac{GMm}{r^2}$$

$$F = \frac{(6.67\times10^{-11})(5.97\times10^{24}\text{ kg})(5.00\text{ kg})}{(6.38\times10^{6}\text{ m})^2}$$

$F = 48.9$ N

13.4 - Shell theorem: inside the sphere

Another section discussed how to calculate the force of gravity exerted on an object on the surface of a sphere (a groundhog on the Earth). Now imagine a groundhog burrowing halfway to the center of the Earth, as shown to the right. For the purposes of calculating the force of gravity, what is the distance between the groundhog and the Earth? And what mass should be used for the Earth in the equation?

The first question is easier: The distance used to calculate gravity's force remains the distance between the groundhog and the Earth's center. Determining what mass to use is trickier. We use the sphere defined by the groundhog's position, as shown to the right. The mass inside this new sphere, and the mass of the groundhog, are used to calculate the gravitational force. (Again, we assume that the Earth's mass is symmetrically distributed.)

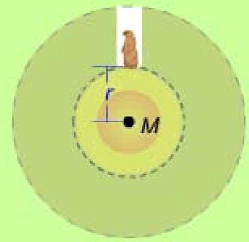
If the groundhog is 10 meters from the Earth's center, the mass enclosed in a sphere with a radius of 10 meters is used in Newton's equation. If the animal moves to the center of the Earth, then the radius of the sphere is zero. At the center, no mass is enclosed, meaning there is no net force of gravity. The groundhog is perhaps feeling a little claustrophobic and warm, but is effectively weightless at the Earth's center.

The volume of a sphere is proportional to the cube of the radius, as the equation to the right shows. If the groundhog burrows halfway to the center of the Earth, then the sphere encloses one-eighth the volume of the Earth and one-eighth the Earth's mass.

Let's place the groundhog at the Earth's center and have him burrow back to the planet's surface. The gravitational force on him increases linearly as he moves back out to the Earth's surface. Why? The force increases proportionally to the mass enclosed by the sphere, which means it increases as the cube of his distance from the center. But the force also decreases as the square of the distance. When the cube of a quantity is divided by its square, the result is a linear relationship.

If the groundhog moves back to the Earth's surface and then somehow moves above the surface (perhaps he boards a plane and flies to an altitude of 10,000 meters), the force again is inversely proportional to the square of the groundhog's distance from the Earth's center. Since the mass of the sphere defined by his position no longer varies, the force is computed using the full mass of the Earth, the mass of the groundhog, and the distance between their centers.

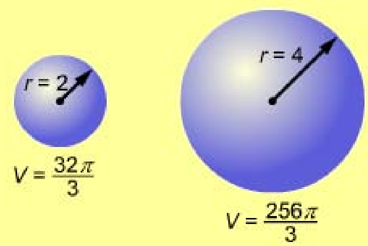
concept 1



Inside a sphere

To calculate gravitational force
- Use mass inside the new shell
- r is distance between object, sphere's center

equation 1



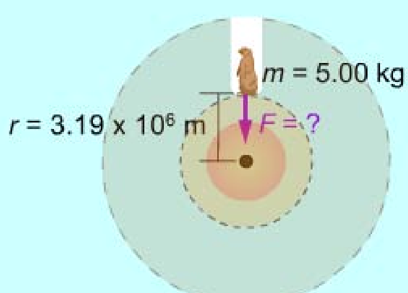
Volume of a sphere

$$V = \frac{4}{3}\pi r^3$$

V = volume

r = radius

13.5 - Sample problem: gravitational force inside the Earth



What is the gravitational force on the groundhog after it has burrowed halfway to the center of the Earth?

Assume the Earth's density is uniform. The Earth's mass and radius are given in the table of variables below.

Variables

gravitational force	F
mass of Earth	$M_E = 5.97\times10^{24}$ kg
radius of Earth	$r_E = 6.38\times10^6$ m
mass of inner sphere	M_s
distance between groundhog and Earth's center	$r_s = 3.19\times10^6$ m
mass of groundhog	$m = 5.00$ kg
gravitational constant	$G = 6.67\times10^{-11}$ N·m²/kg²

What is the strategy?

1. Use the shell theorem. Compute the ratio of the volume of the Earth to the volume of the sphere defined by the groundhog's current position. Use this comparison to calculate the mass of the inside sphere.
2. Use the mass of the inside sphere, the mass of the groundhog and the distance between the groundhog and the center of the Earth to calculate the gravitational force.

Physics principles and equations

Newton's law of gravitation

$$F = \frac{GMm}{r^2}$$

Mathematics principles

The equation for the volume of a sphere is

$$V = \tfrac{4}{3}\pi r^3$$

Step-by-step solution

First, we calculate the mass enclosed by the sphere.

Step	Reason
1. $\dfrac{M_s}{M_E} = \dfrac{V_s}{V_E}$	ratio of masses proportional to ratio of volumes
2. $M_s = \dfrac{V_s M_E}{V_E}$	rearrange
3. $M_s = \dfrac{\left(\frac{4}{3}\pi r_s^3\right)(M_E)}{\frac{4}{3}\pi r_E^3}$	volume of a sphere
4. $M_s = \dfrac{(r_s^3)(M_E)}{r_E^3}$	cancel common factors
5. $M_s = \dfrac{(3.19\times10^6\ \text{m})^3(5.97\times10^{24}\ \text{kg})}{(6.38\times10^6\ \text{m})^3}$	enter values
6. $M_s = 7.46\times10^{23}$ kg	evaluate

Now that we know the mass of the inner sphere, the shell theorem states that we can use it to calculate the gravitational force on the groundhog using Newton's law of gravitation.

Step	Reason
7. $F = \dfrac{GMm}{r^2}$	law of gravitation
8. $F = \dfrac{(6.67\times10^{-11}\frac{\text{N}\cdot\text{m}^2}{\text{kg}^2})(7.46\times10^{23}\text{kg})(5.00\text{kg})}{(3.19\times10^{6}\text{m})^2}$	enter values
9. $F = 24.4\text{ N}$	solve for the force

This calculation also confirms a point made previously: Inside the Earth, the groundhog's weight increases linearly with distance from the planet's center. At half the distance from the center, his weight is half his weight at the surface.

13.6 - Earth's composition and g

The force of gravity differs slightly at different locations on the Earth. These variations mean that g, the acceleration due to the force of gravity, also differs by location. Why do the force of gravity and g vary?

First, the surface of the Earth can be slightly below sea level (in Death Valley, for example), and it can be more than 8000 meters above it (on peaks such as Everest and K2). Compared to an object at sea level, an object at the summit of Everest is 0.14% farther from the center of the Earth. This greater distance to the Earth's center means less force (Newton's law of gravitation), which in turn reduces acceleration (Newton's second law). If you summit Mt. Everest and jump with joy, the force of gravity will accelerate you toward the ground about 0.03 m/s^2 slower than if you were jumping at sea level.

Second, the Earth has a paunch of sorts: It bulges at the equator and slims down at the poles, making its radius at the equator about 21 km greater than at the poles. This is shown in an exaggerated form in the illustration for Concept 3. The bulge is caused by the Earth's rotation and the fact that it is not entirely rigid. This bulge means that at the equator, an object is farther from the Earth's center than it would be at the poles and, again, greater distance means less force and less acceleration.

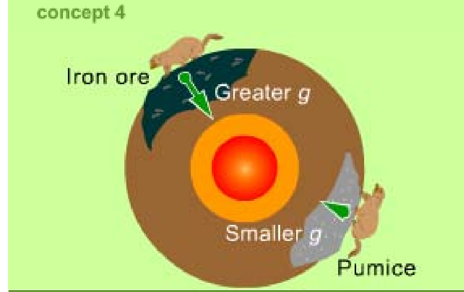
Finally, the density of the planet also varies. The Earth consists of a jumble of rocks, minerals, metals and water. It is denser in some regions than in others. The presence of materials such as iron that are denser than the average increases the local gravitational force by a slight amount. Geologists use *gravity gradiometers* to measure the Earth's density. Variations in the density can be used to prospect for oil or to analyze seismic faults.

Value of *g*
about 9.80 m/s^2 at sea level

Value for *g* varies:
Due to altitude

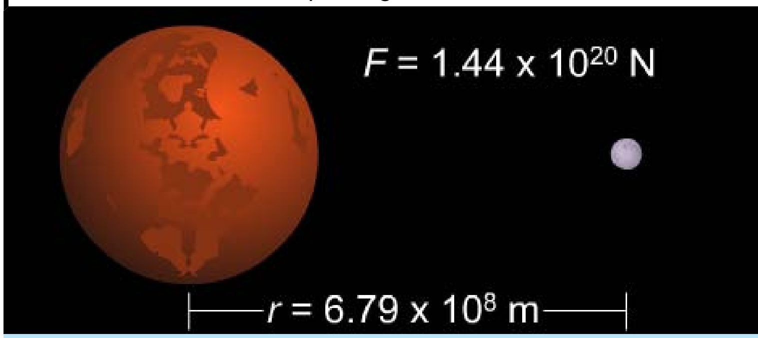
Value for *g* varies:
Because the Earth is not a perfect sphere

concept 4



Value for *g* varies:
Because the Earth is not uniformly dense

13.7 - Interactive checkpoint: gravitation



A planet and its moon are attracted to each other by a gravitational force of 1.44×10^{20} N at a distance of 6.79×10^{8} m. The planet is 155 times as massive as the moon. What is the mass of the planet?

Answer:

$M =$ ______ kg

13.8 - Newton's cannon

In addition to noting that the Earth exerts a force on an apple, Newton also pondered why the Moon circles the Earth. He posed a fundamental question: Is the force the Earth exerts on the Moon the same type of force that it exerts on an apple? He answered yes, and his correct answer would forever change humanity's understanding of the universe.

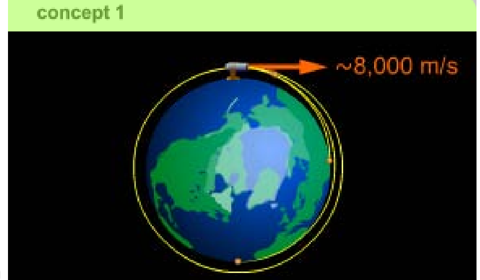
Comparing the orbit of the distant Moon to the fall of a nearby apple required great intellectual courage. Although the motion of the Moon overhead and the fall of the apple may not seem to resemble one another, Newton concluded that the same force dictates the motion of both, leading him to propose new ways to think about the Moon's orbit.

To explain orbital motion, Newton conducted a thought experiment: What would happen if you used a very powerful cannon to fire a stone from the top of a very tall mountain? He knew the stone would obey the basic precepts of projectile motion, as shown in the diagrams to the right.

But, if the stone were fired fast enough, could it just keep going, never touching the ground? (Factors such as air resistance, the Earth's rotation, and other mountains that might block the stone are ignored in Newton's thought experiment.)

Newton concluded that the stone would not return to the Earth if fired fast enough. As he wrote in his work, *Principia*, published in 1686:

concept 1



Newton's cannon
Newton imagined a powerful cannon
The faster the projectile, the farther it travels
At ~8,000 m/s, projectile never touches ground

> " ... the greater the velocity with which [a stone] is projected, the farther it goes before it falls to the earth. We may therefore suppose the velocity to be so increased, that it would describe an arc of 1, 2, 5, 10, 100, 1000 miles before it arrived at earth, till at last, exceeding the limits of the earth, it should pass into space without touching."

Newton correctly theorized that objects in orbit − moons, planets and, today, artificial satellites − are in effect projectiles that are falling around a central body but moving fast enough that they never strike the ground. He could use his theory of gravity and his knowledge of circular motion to explain orbits. (In this section, we focus exclusively on circular orbits, although orbits can be elliptical, as well.)

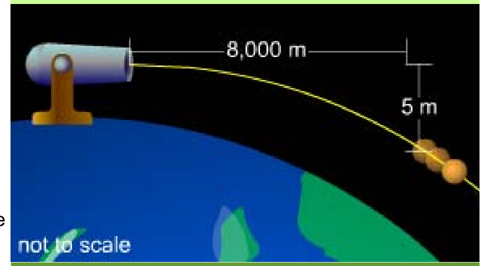
Why is it that the stone does not return to the Earth when it is fired fast enough? Why can it remain in orbit, forever circling the Earth, as shown to the right?

First, consider what happens when a cannon fires a cannonball horizontally from a mountain at a relatively slow speed, say 100 m/s. In the vertical direction, the cannonball accelerates at *g* toward the ground. In the horizontal direction, the ball continues to move at 100 m/s until it hits the ground. The force of gravity pulls the ball down, but there is neither a force nor a change in speed in the horizontal direction (assuming no air resistance).

Now imagine that the cannonball is fired much faster. If the Earth were flat, at some point the ball would collide with the ground. But the Earth is a sphere. Its approximate curvature is such that it loses five meters for every 8000 horizontal meters, as shown in Concept 2. At the proper horizontal (or more properly, tangential) velocity, the cannonball moves in an endless circle around the planet. For every 8000 meters it moves forward, it falls 5 meters due to gravity, resulting in a circle that wraps around the globe.

In this way, satellites in orbit actually are falling around the Earth. The reason astronauts in a space shuttle orbiting close to the Earth can float about the cabin is **not** because gravity is no longer acting on them (the Earth exerts a force of gravity on them), but rather because they are projectiles in freefall.

concept 2



Close-up of Newton's cannon
At high speeds, Earth's curvature affects whether projectile lands
At ~8,000 m/s, ground curves away at same rate that object falls

concept 3

Objects in orbit
Move fast enough to never hit the ground
Continually fall toward the ground, pulled by force of gravity

13.9 - Interactive problem: Newton's cannon

Imagine that you are Isaac Newton. It is the year 1680, and you are staring up at the heavens. You see the Moon passing overhead.

You think: Perhaps the motion of the Moon is related to the motion of Earth-bound objects, such as projectiles. Suppose you threw a stone very, very fast. Is there a speed at which the stone, instead of falling back to the Earth, would instead circle the planet, passing around it in orbit like the Moon?

You devise an experiment in your head, a type of experiment called a thought experiment. A thought experiment is a way physicists can test or explain valuable concepts even though they cannot actually perform the experiment. You ask: "What if I had an extremely powerful cannon mounted atop a mountain. Could I fire a stone so fast it would never hit the ground?"

Try Newton's cannon in the simulation to the right. You control the initial speed of the cannonball by clicking the up and down buttons in the control panel. The cannon fires horizontally, tangent to the surface of the Earth. See if you can put the stone into orbit around the Earth. You can create a nearly perfect circular orbit, as well as orbits that are elliptical.

interactive 1

Fire Newton's cannon to put cannonball into orbit

(In this simulation we ignore the rotation of the Earth, as Newton did in his thought experiment. When an actual satellite is launched, it is fired in the same direction as the Earth's rotation to take advantage of the tangential velocity provided by the spinning Earth.)

13.10 - Circular orbits

The Moon orbits the Earth, the Earth orbits the Sun, and today artificial satellites are propelled into space and orbit above the Earth's surface. (We will use the term *satellite* for any body that orbits another body.)

These satellites move at great speeds. The Earth's orbital speed around the Sun averages about 30,000 m/s (that is about 67,000 miles per hour!) A communications satellite in circular orbit 250 km above the surface of the Earth moves at 7800 m/s.

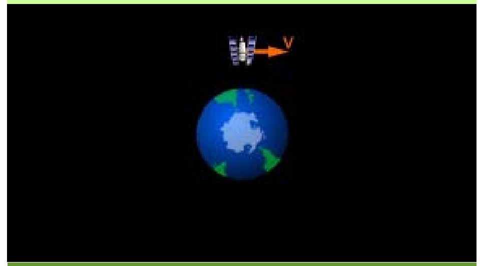
In this section, we focus on circular orbits. Most planets orbit the Sun in roughly circular paths, and artificial satellites typically travel in circular orbits around the Earth as well.

The force of gravity is the centripetal force that along with a tangential velocity keeps the body moving in a circle. We use an equation for centripetal force on the right to derive the relationship between the mass of the body being orbited, orbital radius, and satellite speed. As shown in Equation 1, we first set the centripetal force equal to the gravitational force and then we solve for speed.

This equation has an interesting implication: The mass of the satellite has no effect on its orbital speed. The speed of an object in a circular orbit around a body with mass M is determined solely by the orbital radius, since M and G are constant. Satellite speed and radius are linked in circular orbits. A satellite cannot increase or decrease its speed and stay in the same circular orbit. A change in speed **must** result in a change in orbital radius, and vice versa.

At the same orbital radius, the speed of a satellite increases with the square root of the mass of the body being orbited. A satellite in a circular orbit around Jupiter would have to move much faster than it would if it were in an orbit of the same radius around the Earth.

concept 1

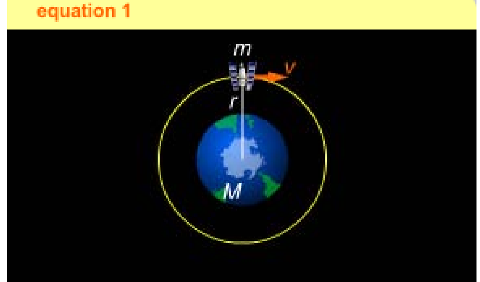


Circular orbits
Satellites in circular orbit have constant speed

concept 2

Orbital speed and radius
Satellite speed and radius are linked
· The smaller the orbit, the greater the speed

equation 1



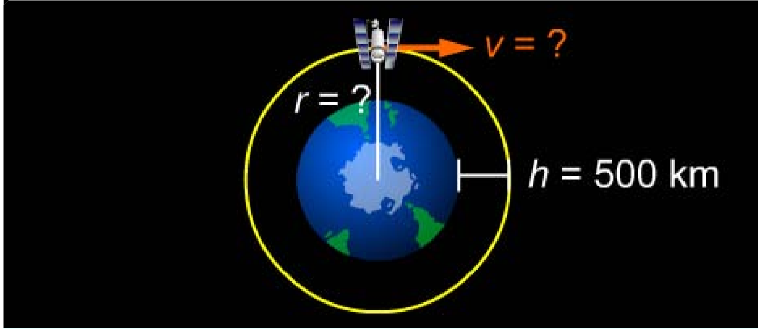
Speed in circular orbit

$$\frac{mv^2}{r} = \frac{GmM}{r^2}$$

$$v = \sqrt{\frac{GM}{r}}$$

m = mass of satellite
v = satellite speed
G = gravitational constant
M = mass of body being orbited
r = orbital radius

13.11 - Sample problem: speed of an orbiting satellite



What is the speed of a satellite in circular orbit 500 kilometers above the Earth's surface?

The illustration above shows a satellite in circular orbit 500 km above the Earth's surface. The Earth's radius is stated in the variables table.

Variables

satellite speed	v
satellite orbital radius	r
satellite height	h = 500 km
Earth's radius	$r_E = 6.38\times10^6$ m
Earth's mass	$M_E = 5.97\times10^{24}$ kg
gravitational constant	$G = 6.67\times10^{-11}$ N·m²/kg²

What is the strategy?

1. Determine the satellite's orbital radius, which is its distance from the center of the body being orbited.
2. Use the orbital speed equation to determine the satellite's speed.

Physics principles and equations

The equation for orbital speed is

$$v = \sqrt{\frac{GM}{r}}$$

Step-by-step solution

We start by determining the satellite's orbital radius. Careful: this is not the satellite's height above the surface of the Earth, but its distance from the center of the planet.

Step	Reason
1. $r = r_E + h$	equation for satellite's orbital radius
2. $r = 6.38\times10^6\text{ m} + 5.00\times10^5\text{ m}$	enter values
3. $r = 6.88\times10^6$ m	add

Now we apply the equation for orbital speed.

Step	Reason
4. $v = \sqrt{\frac{GM_E}{r}}$	orbital speed equation
5. $v = \sqrt{\frac{(6.67\times10^{-11}\frac{\text{N}\cdot\text{m}^2}{\text{kg}^2})(5.97\times10^{24}\text{kg})}{6.88\times10^6\text{m}}}$	enter values
6. v = 7610 m/s	evaluate

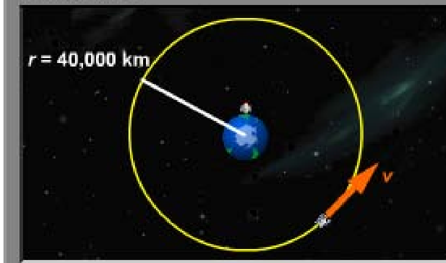
13.12 - Interactive problem: intercept the orbiting satellite

In this simulation, your mission is to send up a rocket to intercept a rogue satellite broadcasting endless *Barney*® reruns. You can accomplish this by putting the rocket into an orbit with the same radius and speed as that of the satellite, but traveling in the opposite direction. The resulting collision will destroy the satellite (and your rocket, but that is a small price to pay to save the world's sanity).

The rogue satellite is moving in a counterclockwise circular orbit 40,000 kilometers (4.00×10^7 m) above the center of the Earth. Your rocket will automatically move to the same radius and will move in the correct direction.

You must do a calculation to determine the proper speed to achieve a circular orbit at that radius. You will need to know the mass of the Earth, which is 5.97×10^{24} kg. Enter the speed (to the nearest 10 m/s) in the control panel and press GO. Your rocket will rise from the surface of the Earth to the same orbital radius as the satellite, and then go into orbit with the speed specified. You do not have to worry about how the rocket gets to the orbit; you just need to set the speed once the rocket is at the radius of the satellite. If you fail to destroy the satellite, check your calculations, press RESET and try a new value.

interactive 1



Send rocket into circular orbit to intercept satellite

13.13 - Interactive problem: dock with an orbiting space station

In this simulation, you are the pilot of an orbiting spacecraft, and your mission is to dock with a space station. As shown in the diagram to the right, your ship is initially orbiting in the same circular orbit as the space station. However, it is on the far side of the Earth from the space station.

In order to dock, your ship must be in the same orbit as the space station, and it must touch the space station. To dock, your speed and radius must be very close to that of the space station. A high speed collision does not equate to docking!

You have two buttons to control your ship. The “Forward thrust” button fires rockets out the back of the ship, accelerating your spacecraft in the direction of its current motion. The other button, labeled “Reverse thrust,” fires retrorockets in the opposite direction. To use more “professional” terms, the forward thrust is called *prograde* and the reverse thrust is called *retrograde*.

Using these two controls, can you figure out how to dock with the space station?

To assist in your efforts, the current orbital paths for both ships are drawn on the screen. Your rocket’s path is drawn in yellow, and the space station’s is drawn in white.

interactive 1

Change speed of blue ship to dock with space station

This simulation requires no direct mathematical calculations, but some thought and experimentation are necessary. If you get too far off track, you may want to press RESET and try again from the beginning.

There are a few things you may find worth observing. You will learn more about them when you study orbital energy. First, what is the change in the speed of the rocket after firing its rear (Forward thrust) engine? What happens to its speed after a few moments? If you **qualitatively** consider the total energy of the ship, can you explain you what is going on? You may want to consider it akin to what happens if you throw a ball straight up into the sky.

If you cannot dock but are able to leave and return to the initial circular orbit, you can consider your mission achieved.

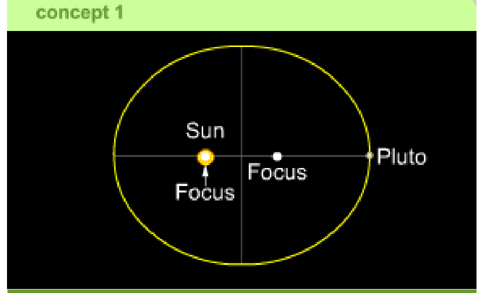
13.14 - Kepler’s first law

The law of orbits: Planets move in elliptical orbits around the Sun.

The reason Newton’s comparison of the Moon’s motion to the motion of an apple was so surprising was that many in his era believed the orbits of the planets and stars were “divine circles:” arcs across the cosmic sky that defied scientific explanation. Newton used the fact that the force of gravity decreases with the square of the distance to explain the geometry of orbits.

Scientists had been proposing theories about the nature of orbits for centuries before Newton stated his law of gravitation. Numerous theories held that all bodies circle the Earth, but subsequent observations began to point to the truth: the Earth and other planets orbit the Sun. The conclusion was controversial; in 1633 the Catholic Church forced Galileo to repudiate his writings that implied this conclusion.

concept 1



Kepler’s first law
Planets move in elliptical orbits with the Sun at one focus

Even earlier, in 1609, the astronomer and astrologer Johannes Kepler began to propose what are now three basic laws of astronomy. He developed these laws through careful mathematical analysis, relying on the detailed observations of his mentor, Tycho Brahe, a talented and committed observational astronomer.

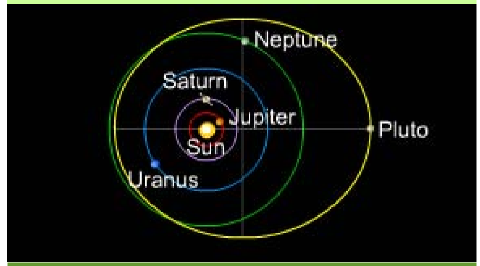
Kepler and Brahe formed one of the most productive teams in the history of astronomy. Brahe had constructed a state of the art observatory on an island off the coast of Denmark. “State of the art” is always a relative term − the telescope had not yet been invented, and Brahe might well have traded his large observatory for a good pair of current day binoculars. However, Brahe's records of years of observations allowed Kepler, with his keen mathematical insight, to derive the fundamental laws of planetary motion. He accomplished this decades before Newton published his law of gravitation.

Kepler, using Brahe's observations of Mars, demonstrated what is now known as Kepler's first law. This law states that all the planets move in elliptical orbits, with the Sun at one focus of the ellipse.

The planets in the solar system all move in elliptical motion. The distinctly elliptical orbit of Pluto is shown to the right, with the Sun located at one focus of the ellipse. Had Kepler been able to observe Pluto, the elliptical nature of orbits would have been more obvious.

The other planets in the solar system, some of which he could see, have orbits that are very close to circular. (If any of them moved in a perfectly circular orbit, they would still be moving in an ellipse, since a circle is an ellipse with both foci at its center.) Some of the orbits of these other planets are shown in Concept 2.

concept 2



Solar system
Most planetary orbits are nearly circular

13.15 - More on ellipses and orbits

The ellipse shape is fundamental to orbits and can be described by two quantities: the semimajor axis a and the eccentricity e. Understanding these properties of an ellipse proves useful in the study of elliptical orbits.

The *semimajor axis,* represented by $a,$ is one-half the width of the ellipse at its widest, as shown in Concept 1. You can calculate the semimajor axis by averaging the maximum and minimum orbital radii, as shown in Equation 1.

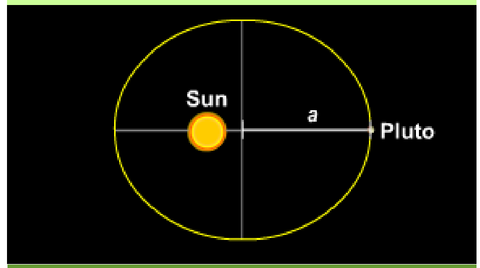
The *eccentricity* is a measure of the elongation of an ellipse, or how much it deviates from being circular. (The word eccentric comes from "ex-centric," or off-center.) Mathematically, it is the ratio of the distance d between the ellipse's center and one focus to the length a of its semimajor axis. You can see both these lengths in Equation 2. Since a circle's foci are at its center, d for a circle equals zero, which means its eccentricity equals zero.

Pluto has the most eccentric orbit in our solar system, with an eccentricity of 0.25, as calculated on the right. By comparison, the eccentricity of the Earth's orbit is 0.0167. Most of the planets in our solar system have nearly circular orbits.

Comets have extremely eccentric orbits. This means their distance from the Sun at the *aphelion*, the point when they are farthest away, is much larger than their distance at the *perihelion*, the point when they are closest to the Sun. (Both terms come from the Greek, "far from the Sun" and "near the Sun" respectively.)

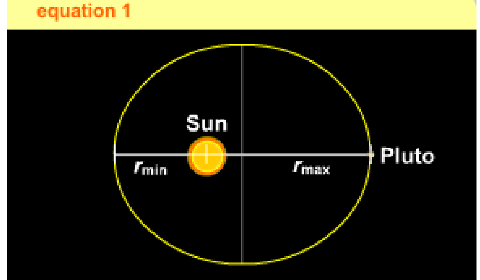
Halley's comet has an eccentricity of 0.97. We show the comet's orbit in Example 1, but for visual clarity it is not drawn to scale. The comet's orbit is so eccentric that its maximum distance from the Sun is 70 times greater than its minimum distance. In Example 1, you calculate the perihelion of this object in *AU*. The AU (*astronomical unit*) is a unit of measurement used in planetary astronomy. It is equal to the average radius of the Earth's orbit around the Sun: about $1.50{\times}10^{11}$ meters.

concept 1



Elliptical orbits
Semimajor axis: one half width of orbit at widest
Eccentricity: elongation of orbit

equation 1

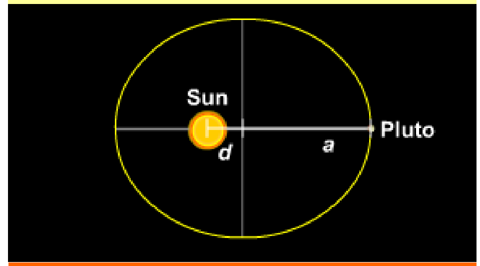


Semimajor axis

$$a = \frac{r_{min} + r_{max}}{2}$$

a = semimajor axis
r_{min} = minimum orbital distance
r_{max} = maximum orbital distance

equation 2

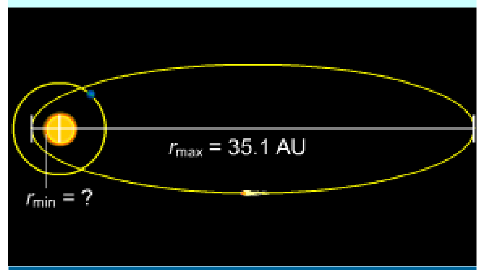


Eccentricity

$$e = \frac{d}{a}$$

e = eccentricity
d = distance from center to focus
a = semimajor axis

example 1

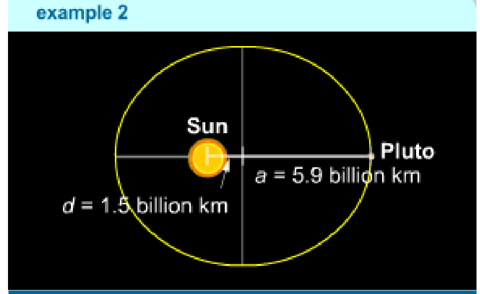


What is the perihelion distance of Halley's Comet? The comet's semimajor axis is 17.8 AU.

$$a = \frac{r_{min} + r_{max}}{2}$$

$r_{min} = 2a - r_{max}$
$r_{min} = 2(17.8 \text{ AU}) - 35.1 \text{ AU}$
$r_{min} = 0.5 \text{ AU}$

example 2



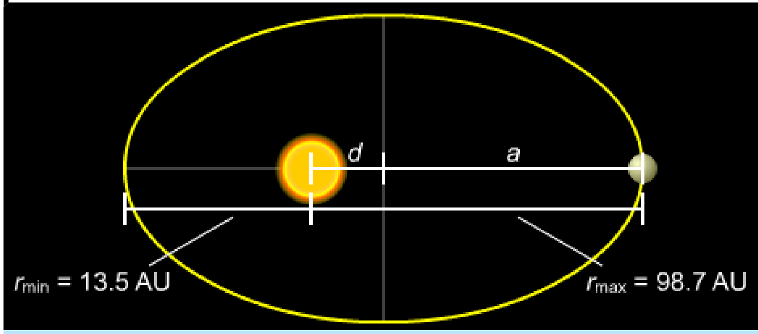
What is the eccentricity of Pluto's orbit?

$$e = \frac{d}{a}$$

$$e = \frac{1.5 \text{ billion km}}{5.9 \text{ billion km}}$$

$e = 0.25$

13.16 - Interactive checkpoint: elliptical orbit



An asteroid orbits a star in an elliptical orbit with a periapsis (closest approach) of 13.5 AU and an apoapsis (farthest distance) of 98.7 AU. What are the semimajor axis and the eccentricity of the asteroid's orbit?

Answer:

a = ______ AU

e = ______

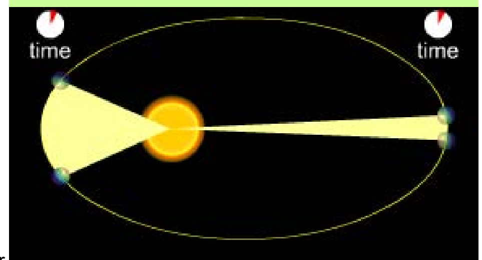
13.17 - Kepler's second law

The law of areas: An orbiting body sweeps out equal areas in equal amounts of time.

In his second law, Kepler used a geometrical technique to show that the speed of an orbiting planet is related to its distance from the Sun. (We use the example of a planet and the Sun; this law applies equally well for a satellite orbiting the Earth, or for Halley's comet orbiting the Sun.)

Kepler used the concept of a line connecting the planet to the Sun, moving like a second hand on a watch. As shown to the right, the line "sweeps out" slices of area over time. His second law states that the planet sweeps out an equal area in an equal amount of time in any part of an orbit. In an elliptical orbit, planets move slowest when they are farthest from the Sun and move fastest when they are closest to the Sun.

concept 1



Kepler's second law
Planets in orbit sweep out equal areas in equal times

Kepler established his second law nearly a century before Newton proposed his theory of gravitation. Although Kepler did not know that gravity varied with the inverse square of the distance, using Brahe's data and his own keen quantitative insights he determined a key aspect of elliptical planetary motion.

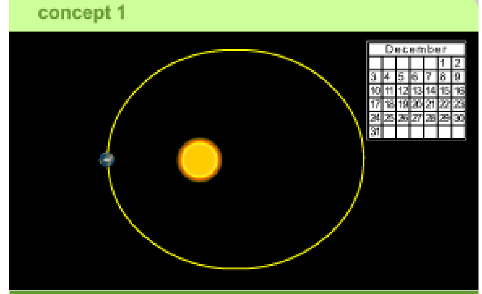
13.18 - Kepler's third law

The law of periods: The square of the period of an orbit is proportional to the cube of the semimajor axis of the orbit.

Kepler's third law, proposed in 1619, states that the period of an orbit around a central body is a function of the semimajor axis of the orbit and the mass of the central body. The semimajor axis a is one half the width of the orbit at its widest. In a circular orbit, the semimajor axis is the same as the radius r of the orbit.

We illustrate this in Concept 1 using the Earth and the Sun. Given the scale of illustrations in this section, the Earth's nearly circular orbit appears as a circle.

concept 1



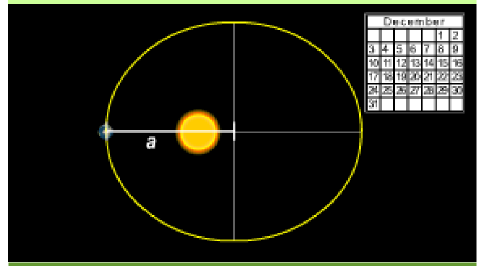
Orbital period
Time of one revolution

The length of the Earth's period − a year, the time required to complete a revolution about the Sun − is solely a function of the mass of the Sun and the distance a shown in Concept 2.

Kepler's third law states that the square of the period is proportional to the cube of the semimajor axis, and inversely proportional to the mass of the central body. The law is shown in Equation 1. For the equation to hold true, the mass of the central body must be much greater than that of the satellite.

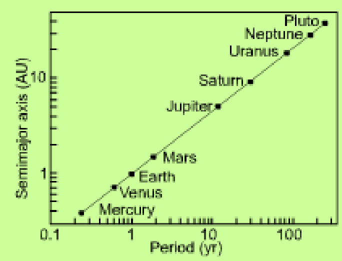
This law has an interesting implication: The square of the period divided by the cube of the semimajor axis has the same value for all the bodies orbiting the Sun. In our solar system, that ratio equals about 3×10^{-34} years^2/meters^3 (where "years" are Earth years) or 3×10^{-19} s^2/m^3. This is demonstrated in the graph in Concept 3. The horizontal and vertical scales of the coordinate system are logarithmic, with semimajor axis measured in AU and period measured in Earth years.

concept 2



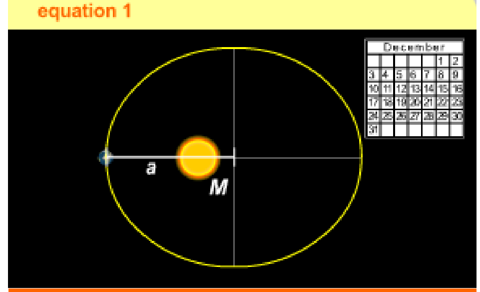
Kepler's third law
Square of orbital period proportional to cube of semimajor axis

concept 3



Graph of Kepler's third law
Orbital size versus period for planets orbiting Sun

equation 1

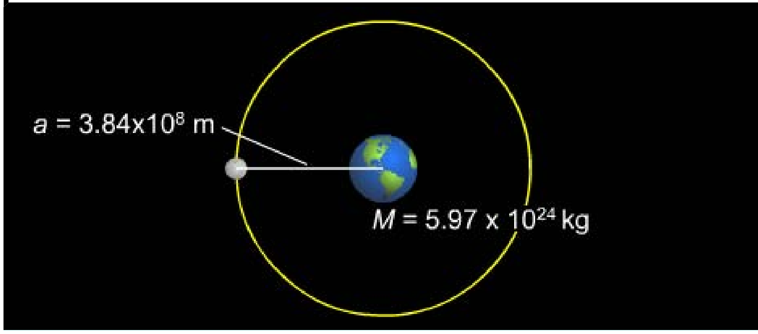


Kepler's third law

$$T^2 = \left(\frac{4\pi^2}{GM}\right)a^3$$

T = satellite period (in seconds)
G = gravitational constant
M = mass of central body
a = semimajor axis

13.19 - Sample problem: the period of the Moon



What is the period of the Moon's orbit, in days?

The illustration shows the Moon in its slightly elliptical orbit around the Earth, with a semimajor axis of 3.84×10^8 m. In this problem, we calculate the *sidereal period*, the time it takes for the Moon to return to the same position in the sky, using a background of "fixed" stars as a reference. The Moon's *synodic period*, the time it takes to return to the same position relative to the Sun, is longer.

Variables

Moon period	T
Moon's orbital semimajor axis	$a = 3.84\times10^8$ m
Earth's mass	$M_E = 5.97\times10^{24}$ kg
gravitational constant	$G = 6.67\times10^{-11}$ N·m²/kg²

What is the strategy?

Use Kepler's third law to find the Moon's orbital period in seconds and convert the result to days.

Physics principles and equations

We will use Kepler's third law

$$T^2 = \left(\frac{4\pi^2}{GM}\right)a^3$$

Step-by-step solution

Enter the given values into the equation for Kepler's third law. Be careful with units: The equation for Kepler's third law gives the period in seconds.

Step	Reason
1. $T^2 = \left(\frac{4\pi^2}{GM}\right)a^3$	Kepler's third law
2. $T^2 = \frac{4\pi^2(3.84\times10^8\ \text{m})^3}{(6.67\times10^{-11}\ \frac{\text{N}\cdot\text{m}^2}{\text{kg}^2})(5.97\times10^{24}\ \text{kg})}$	enter values
3. $T = 2.37\times10^6$ s	evaluate
4. $T \text{ in days} = (2.37\times10^6\ \text{s})\frac{(1\ \text{day})}{(8.64\times10^4\ \text{s})}$	convert seconds to days
5. T in days = 27.4 days	evaluate

13.20 - Interactive problem: geosynchronous satellite

In this simulation, your mission is to put a satellite into circular geosynchronous orbit around the Earth.

A geosynchronous orbit is one in which the satellite's orbital period is 24 hours, the same as the Earth's rotational period. If situated over the equator, a geosynchronous satellite stays positioned over the same location, a requirement for some communications satellites.

What orbital radius and orbital speed must you set for the satellite so that it achieves a circular, geosynchronous orbit? You need to determine a speed that will cause the satellite to remain over a fixed point above the Earth. (Hint: There are 24 hours in a day….) Kepler's third law will help you determine the orbital radius.

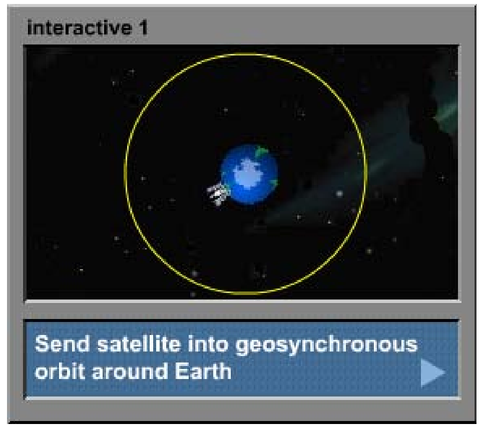
Once you have determined the required radius, use that value to determine the orbital speed for a satellite in a circular orbit.

This simulation contains an added complication: You need to factor in the Earth's motion. The Earth supplies a tangential velocity of 460 m/s to satellites launched from the equator, and the rocket launches in the direction that allows it to take advantage of that “boost”. In short, you must subtract 460 m/s from whatever you have determined is the orbital speed needed for geosynchronous orbit. The ship will keep that speed and maneuver appropriately as it moves to the radius orbit you specified.

Enter an orbital radius (in kilometers) and a launch speed (in m/s) for your satellite in the entry boxes in the control panel. When you press GO, the satellite will launch according to your specifications.

If the satellite lines up with the rotating red dot on Earth, you have succeeded. (A yellow circle indicates the size of the necessary orbit.) If you do not succeed at first, press RESET, redo your calculations, and blast off again.

Note that we are showing a view of the Earth from above the South Pole, and from that view the Earth rotates in a clockwise direction.



13.21 - Orbits and energy

Satellites have both kinetic energy and potential energy. The KE and PE of a satellite in elliptical orbit both change as it moves around its orbit. This is shown in Concept 1. Energy gauges track the satellite’s changing PE and KE. When the satellite is closer to the body it orbits, Kepler’s second law states that it moves faster, and greater speed means greater KE.

The PE of the system is less when the satellite is closer to the body it orbits. When discussing gravitational potential energy, we must choose a reference point that has zero PE. For orbiting bodies, that reference point is usually defined as infinite separation. As two bodies approach each other from infinity, potential energy decreases and becomes increasingly negative as the value declines from zero.

Concept 2 shows that even while the PE and KE change continuously in an elliptical orbit, the total energy TE stays constant. Because there are no external forces acting on the system consisting of the satellite and the body it orbits, nor any internal dissipative forces, its total mechanical energy must be conserved. Any increase in kinetic energy is matched by an equivalent loss in potential energy, and vice versa.

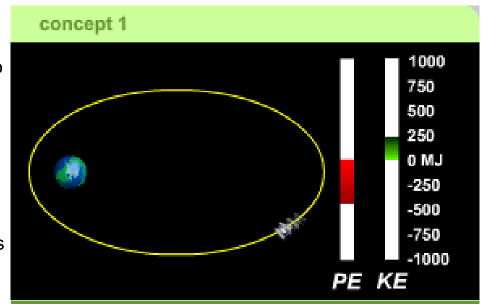
In a circular orbit, a satellite’s speed is constant and its distance from the central body remains the same, as shown in Concept 3. This means that both its kinetic and potential energies are constant.

The total energy of a satellite increases with the radius (in the case of circular orbits) or the semimajor axis (in the case of elliptical orbits). Moving a satellite into a larger orbit requires energy; the source of that energy for a satellite might be the chemical energy present in its rocket fuel.

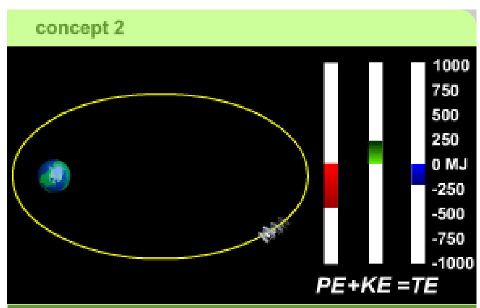
Equations used to determine the potential and kinetic energies and the total energy of a satellite in **circular** orbit are shown in Equation 1. These equations can be used to determine the energy required to boost a satellite from one circular orbit to another with a different radius. The KE equation can be derived from the equation for the velocity of a satellite. The PE equation holds true for any two bodies, and can be derived by calculating the work done by gravity as the satellite moves in from infinity.

The equations have an interesting relationship: The kinetic energy of the satellite equals one-half the absolute value of the potential energy. This means that when the radius of a satellite’s orbit increases, the total energy of the satellite increases. Its kinetic energy decreases since it is moving more slowly in its higher orbit, but the potential energy increases twice as much as this decrease in KE.

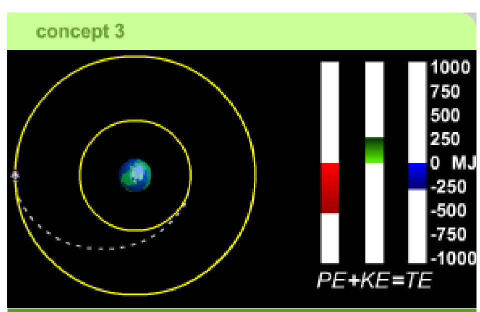
Equation 2 shows the total energy equation for a satellite in an elliptical orbit. This equation uses the value of the semimajor axis a instead of the radius r.



Orbital energy
Satellites have kinetic and potential energy
Since PE = 0 at infinite distance, PE always negative



For a given orbit:
Total energy is constant



For a given circular orbit
Both PE, KE constant
Total energy increases with radius

equation 1

Energy in circular orbits

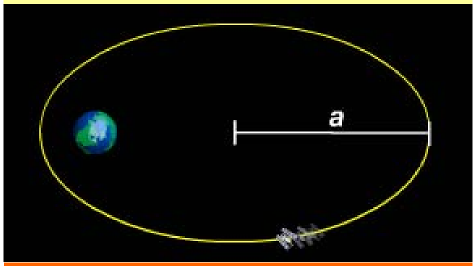
$$KE = \frac{GMm}{2r}$$

$$PE = -\frac{GMm}{r}$$

$$E_{tot} = -\frac{GMm}{2r}$$

G = gravitational constant
M = mass of planet
m = mass of satellite
r = orbit radius

equation 2

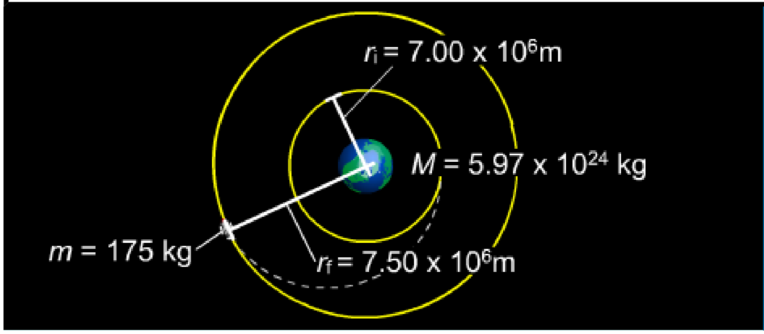


Energy in elliptical orbits

$$E_{tot} = -\frac{GMm}{2a}$$

E_{tot} = total energy
a = semimajor axis

13.22 - Sample problem: energy and orbital radius



A 175 kg satellite in circular orbit fires its rockets twice to move to a new circular orbit with a larger radius. How much energy does the satellite expend to attain the wider orbit?

The illustration shows a satellite in a circular orbit performing a two-stage maneuver called a *Hohmann transfer* that boosts it to a circular orbit with a larger radius. The diagram is not drawn to scale. We assume that the satellite's mass stays effectively constant.

Variables

first orbital radius	$r_i = 7.00\times10^6$ m
second orbital radius	$r_f = 7.50\times10^6$ m
mass of Earth	$M = 5.97\times10^{24}$ kg
mass of satellite	$m = 175$ kg
gravitational constant	$G = 6.67\times10^{-11}$ N·m²/kg²
total energy of satellite	E

What is the strategy?

1. Write an equation for the change in the total energy of the satellite.
2. Substitute the total energy expressions for the satellite in each orbit.
3. Enter the given values and calculate the work performed.

Physics principles and equations

The work done by the satellite's rocket engine equals its change in total energy.

The equation for total energy of a satellite in a circular orbit is,

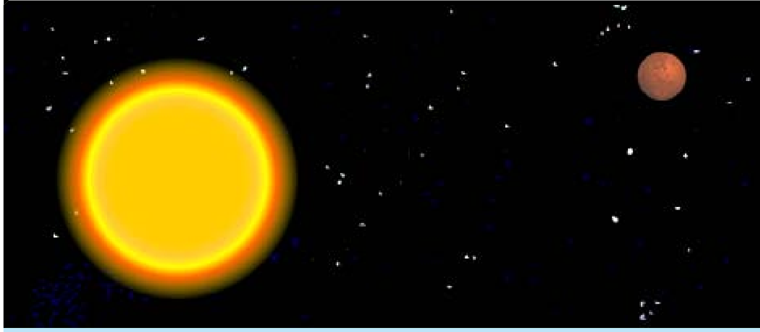
$$E = -\frac{GMm}{2r}$$

Step-by-step solution

The change in the satellite's energy is equal to its final energy minus its initial energy. We write this as an equation, then substitute the energy expressions for the satellite in each orbit and evaluate the result.

Step	Reason
1. $\Delta E = E_f - E_i$	change in energy
2. $\Delta E = \left(-\frac{GMm}{2r_f}\right) - \left(-\frac{GMm}{2r_i}\right)$ $\Delta E = \frac{GMm}{2}\left(\frac{1}{r_i} - \frac{1}{r_f}\right)$	substitute energy expressions
3. $\Delta E = \frac{GMm}{2}\left(\frac{1}{(7.00\times10^6\text{ m})} - \frac{1}{(7.50\times10^6\text{ m})}\right)$ $\Delta E = \frac{GMm}{2(1.05\times10^8\text{ m})}$	substitute values for radii and simplify
4. $\Delta E = \frac{(6.67\times10^{-11}\frac{\text{N}\cdot\text{m}^2}{\text{kg}^2})(5.97\times10^{24}\text{ kg})(175\text{ kg})}{2(1.05\times10^8\text{ m})}$	enter remaining values
5. $\Delta E = 3320\times10^5\text{ J} = 3.32\times10^8\text{ J}$	evaluate change in total energy

13.23 - Interactive checkpoint: Kepler's third law and energy



A planet orbits a star with a period of 7.78×10^{6} s and a semimajor axis of 7.49×10^{10} m. What is the mass of the star?

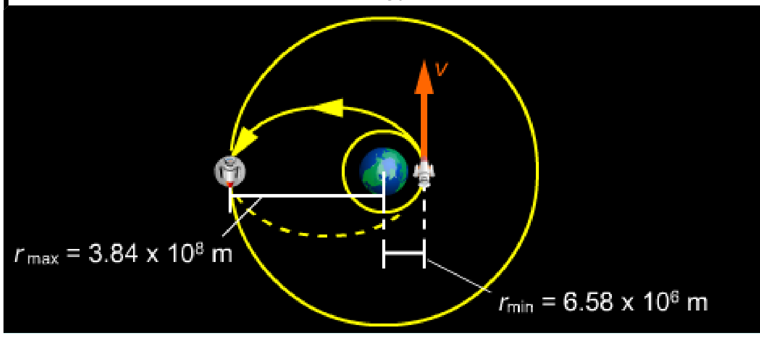
The total energy of the planet in this orbit is -1.04×10^{37} J. What is the mass of the planet?

Answer:

$M =$ ______ kg

$m =$ ______ kg

13.24 - Sample problem: energy of a rocket to the Moon



A 20,000 kg rocket is in a circular orbit above the Earth having a radius $6.58\text{x}10^{6}$ m. How much energy will it take to boost the rocket to the Moon on the orbital path shown?

Upon firing its engines, the rocket will follow an elliptical orbital path to the Moon.

This problem simplifies the actual orbital mechanics required to fire a rocket to the Moon. For instance, the question ignores the gravitational influence of other bodies, such as the Moon and the Sun. However, it provides a good starting point for determining the energy.

Variables

energy of circular orbit	E_c
energy of elliptical orbit	E_e
radius of circular orbit	$r_{min} = 6.58\times10^{6}$ m
Earth-Moon distance	$r_{max} = 3.84\times10^{8}$ m
semimajor axis of ellipse	a
mass of Earth	$M = 5.97\times10^{24}$ kg
mass of rocket	$m = 2.00\times10^{4}$ kg
gravitational constant	$G = 6.67\times10^{-11}$ N·m^2/kg^2

What is the strategy?

1. Determine the energy of the circular orbit around the Earth.
2. Using the dimensions of the elliptical transfer orbit to the Moon, determine the semimajor axis of the elliptical orbit. A diagram useful for accomplishing this is supplied below.
3. Use the semimajor axis to determine the energy of the elliptical transfer orbit.
4. The energy required by the rocket is the energy of the elliptical orbit minus the energy of the circular orbit.

Physics principles and equations

The total energy of a circular orbit is

$$E = -\frac{GMm}{2r}$$

The total energy of an elliptical orbit is

$$E = -\frac{GMm}{2a}$$

The semimajor axis a of an elliptical orbit is the average of its perigee and apogee distances.

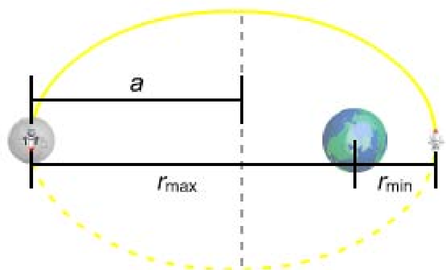
$$a = \frac{r_{min} + r_{max}}{2}$$

Step-by-step solution

First, we calculate the rocket's energy as it orbits the Earth in its initial circular orbit.

Step	Reason
1. $E_c = -\frac{GMm}{2r}$	energy of circular orbit
2. $E_c = -\frac{(6.67\times10^{-11}\frac{\text{N}\cdot\text{m}^2}{\text{kg}^2})(5.97\times10^{24}\text{kg})(2.00\times10^{4}\text{kg})}{2(6.58\times10^{6}\text{m})}$	enter values
3. $E_c = -6.05\times10^{11}$ J	solve for orbital energy

The rocket will be fired along an elliptical path to intersect with the Moon's orbit. We use the dimensions of the elliptical orbit to determine the rocket's change in energy.



The perigee distance of the elliptical orbit equals the radius of the initial circular orbit around the Earth. The apogee distance of the new elliptical orbit equals the Earth-Moon distance. We use these facts below to determine the semimajor axis a of the elliptical orbit.

Step	Reason
4. $a = \frac{r_{min} + r_{max}}{2}$	equation for semimajor axis
5. $a = \frac{(6.58\times10^{6}\text{ m} + 3.84\times10^{8}\text{ m})}{2}$	enter values for perigee and apogee
6. $a = 1.95\times10^{8}$ m	solve for semimajor axis

Now we can determine the energy of the rocket in its elliptical transfer orbit. Note that it is still in orbit around the Earth.

Step	Reason
7. $E_e = -\frac{GMm}{2a}$	energy of elliptical orbit
8. $E_e = -\frac{(6.67\times10^{-11}\frac{\text{N}\cdot\text{m}^2}{\text{kg}^2})(5.97\times10^{24}\text{kg})(2.00\times10^{4}\text{kg})}{2(1.95\times10^{8}\text{m})}$	enter values
9. $E_e = -2.04\times10^{10}$ J	solve for energy

The energy required to boost the rocket from the circular orbit to the elliptical orbit is the difference between the two orbital energies.

Step	Reason
10. $\Delta E = E_e - E_c$	change in energy
11. $\Delta E = -2.04\times10^{10}\ \text{J} - (-6.05\times10^{11}\ \text{J})$	enter values
12. $\Delta E = 5.85\times10^{11}\ \text{J}$	solve for change in energy

13.25 - Interactive problem: a rocket mission to Mars

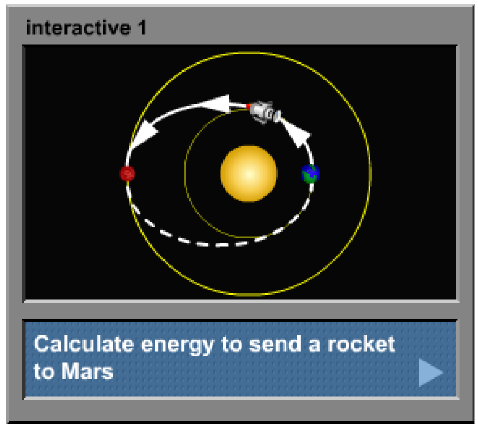
In this simulation, you are a space engineer planning a manned mission to Mars.

You can use what you know about the energy of orbits to calculate how much energy is required to propel a rocket from Earth to Mars. NASA currently is contemplating just such a manned mission to Mars.

Your 100,000-kg rocket is near the Earth. The rocket's mass is treated as constant. The radius of its circular orbit around the Sun is the same as the Earth's. By firing the engines of your rocket, you can move the rocket out of this orbit and into a new elliptical orbit about the Sun. As part of that elliptical orbit, it will rendezvous with Mars.

To make this trip, first calculate the ship's energy due to its circular orbit about the Sun. Then, calculate the energy of the elliptical transfer orbit required to get to Mars.

To calculate the energy of the elliptical transfer orbit to Mars, you must first calculate the semimajor axis of this orbit. The orbit's maximum and minimum distances are the Sun-Mars distance and Sun-Earth distance, respectively. The orbit is tangential to these two orbits. These values are shown below.



The difference in energy between the elliptical transfer orbit to Mars and the ship's initial circular orbit around the Sun is the amount of energy you must enter in the control panel in order to make the trip successfully. You enter the value in joules in the simulation with scientific notation, setting both the leading value (to the nearest tenth) and the exponent.

The table below contains the data you need to solve the problem.

Variables

Earth-Sun distance	$r_{E-S} = 1.50\times10^{11}$ m
Mars-Sun distance	$r_{M-S} = 2.28\times10^{11}$ m
mass of rocket	$m = 1.00\times10^5$ kg
mass of Sun	$M = 1.99\times10^{30}$ kg
gravitational constant	$G = 6.67\times10^{-11}\ \text{N·m}^2/\text{kg}^2$
energy of ship's circular orbit	E_c
semimajor axis of elliptical transfer orbit	a
energy of elliptical transfer orbit	E_e
change in energy	ΔE

In this simulation we ask you to calculate only the energy required to reach Mars. It would take still more energy to enter Mars' orbit and land on the planet. The simulation also ignores the gravitational effects of other planets.

13.26 - Escape speed

Escape speed: The minimum speed required for an object to escape a planet's gravitational attraction.

You know that if you throw a ball up in the air, the gravitational pull of the Earth will cause it to fall back down. If by some superhuman burst of strength you were able to hurl the ball up fast enough, it could have enough speed that the force of Earth's gravity would never be able to slow it down enough to cause it to return to the Earth.

The speed required to accomplish this feat is called the escape speed. Space agencies frequently fire rockets with sufficient speed to escape the Earth's gravity as they explore space.

Given enough speed, a rocket can even escape the Sun's gravitational influence, allowing it to explore outside our solar system. As an example, the Pioneer 10 spacecraft, launched in 1972, was nearly 8 billion miles away from the Earth by 2003, and is projected by NASA to continue to coast silently through deep space into the interstellar reaches indefinitely.

At the right is an equation for calculating the escape speed from a planet of mass M. As the example problem shows, the escape speed for the Earth is about 11,200 m/s, a little more than 40,000 km/h. The escape **speed** does not depend on the mass of the object being launched. However, the **energy** given to the object to make it escape is a function of its mass, since the object's kinetic energy is proportional to its mass.

The rotation of the Earth is used to assist in the gaining of escape speed. The Earth's rotation means that a rocket will have tangential speed (except at the poles, an unlikely launch site for other reasons as well). The tangential speed equals the product of the Earth's angular velocity and the distance from the Earth's axis of rotation.

An object will have a greater tangential speed near the equator because there the distance from the Earth's axis of rotation is greatest. The United States launches its rockets from as close to the equator as is convenient: southern Florida. The rotation of the Earth supplies an initial speed of 1469 km/h (408 m/s) to a rocket fired east from Cape Canaveral, about 4% of the required escape speed.

Derivation. We will derive the escape speed equation by considering a rocket launched from a planet of mass M with initial speed v. The rocket, pulled by the planet's gravity, slows as it rises. Its launch speed is just large enough that it never starts falling back toward the planet; instead, its speed approaches zero as it approaches an infinite distance from the planet. If the initial speed is just a little less, the rocket will eventually fall back toward the planet. If the speed is greater than or equal to the escape speed, the rocket will never return.

Variables

We will derive the escape speed by comparing the potential and kinetic energies of the rocket as it blasts off (subscript 0), and as it approaches infinity (subscript ∞).

	initial	final	change
potential energy	PE_0	PE_∞	ΔPE
kinetic energy	KE_0	KE_∞	ΔKE
gravitational constant	G		
mass of planet	M		
mass of rocket	m		
radius of planet	r		
initial speed of rocket	v		

Strategy

1. Calculate the change in potential energy of the Earth-rocket system between launch and an infinite separation of the two.
2. Calculate the change in kinetic energy between launch and an infinite separation.
3. The conservation of mechanical energy states that the total energy after the engines have ceased firing equals the total energy at infinity. State this relationship and solve for v, the initial escape speed.

Physics principles and equations

We use the equation for the potential energy of an object at a distance r from the center of the planet.

$$PE = -\frac{GMm}{r}$$

We use the definition of kinetic energy.

$$KE = \frac{1}{2}mv^2$$

Finally, we will use the equation that expresses the conservation of mechanical energy.

$$\Delta PE + \Delta KE = 0$$

concept 1

Escape speed
Minimum speed to escape planet's gravitational attraction

concept 2

Body	Mass (kg)	Radius (km)	Escape speed (km/s)
Pluto	1.5×10^{22}	1151	1.1
Earth's Moon	7.35×10^{22}	1738	2.37
Mars	6.42×10^{23}	3397	5.02
Earth	5.97×10^{24}	6378	11.2
Jupiter	1.90×10^{27}	71,492	59.6
Sun	1.99×10^{30}	695,990	618

Some escape speeds

equation 1

Escape speed equation

$$v = \sqrt{\frac{2GM}{r}}$$

v = escape speed
G = gravitational constant
M = mass of planet
r = radius of planet

Step-by-step derivation

In the first steps we find the potential energy of the rocket at launch and at infinity, and subtract the two values.

Step	Reason
1. $PE_0 = -\frac{GMm}{r}$	potential energy at launch
2. $PE_\infty = -\frac{GMm}{\infty} = 0$	potential energy at infinity
3. $\Delta PE = \frac{GMm}{r}$	change in potential energy

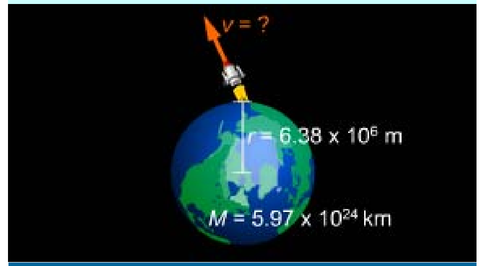
In the following steps we find the kinetic energy of the rocket at launch and at infinity, and subtract the two values.

Step	Reason
4. $KE_0 = \frac{mv^2}{2}$	kinetic energy
5. $KE_\infty = 0$	assumption
6. $\Delta KE = -\frac{mv^2}{2}$	change in kinetic energy

To complete the derivation, we will substitute both changes in energy into the equation for the conservation of energy, and solve for the critical speed v.

Step	Reason
7. $\Delta PE + \Delta KE = 0$	conservation of energy
8. $\frac{GMm}{r} + \left(-\frac{mv^2}{2}\right) = 0$	substitute equations 3 and 6 into equation 7
9. $v = \sqrt{\frac{2GM}{r}}$	solve for v

example 1



What is the minimum speed required to escape the Earth's gravity?

$$v = \sqrt{\frac{2GM}{r}}$$

$$v = \sqrt{\frac{2(6.67\times10^{-11})(5.97\times10^{24}\,\mathrm{kg})}{6.38\times10^{6}\,\mathrm{m}}}$$

$v = 11{,}200$ m/s

13.27 - Gotchas

The freefall acceleration rate, g, does not depend on the mass that is falling. If you say yes, you are agreeing with Galileo and you are correct.

A satellite can move faster and yet stay in the same circular orbit. No, it cannot. The speed is related to the dimensions of the orbit.

13.28 - Summary

Newton's law of gravitation states that the force of gravity between two particles is proportional to the product of their masses and inversely proportional to the square of the distance between them.

The gravitational constant G appears in Newton's law of gravitation and is the same everywhere in the universe. It should not be confused with g, the acceleration due to gravity near the Earth's surface. The value of g varies slightly according to the location on the Earth. This is due to local variations in altitude and to the Earth's bulging shape, nonuniform density and rotation.

The shell theorem states that the force of gravity outside a sphere can be calculated by treating a uniform sphere's mass as though it were all concentrated at its center. For example, to use Newton's law of gravitation to calculate the gravitational force the Earth exerts on an object at its surface, the distance used in the equation is the radius of the Earth.

Inside a sphere, the shell theorem states that the net gravitational force on an object can be calculated by using only the mass located inside the smaller spherical shell defined by the position of the object. The distance is still measured between the object and the center of the sphere.

A gravitational field expresses the gravitational force per unit mass at a given point in space and points in the direction of the force there. If we visualize the gravitational field by using a diagram, field lines are closer together where the field is stronger.

A circular orbit is the simplest type of orbit. The speed of an object in a circular orbit is inversely proportional to the square root of the orbital radius. The speed is also proportional to the square root of the mass of the object being orbited, so orbiting a more massive object requires a greater speed to maintain the same radius.

Johannes Kepler set forth three laws that describe the orbital motion of planets. Kepler's first law says that the planets move in elliptical orbits around the Sun, which is located at one focus of the ellipse. Most of the planets' orbits in the solar system are only slightly elliptical.

An ellipse can be described by two values. The semimajor axis is one half the width of the longest axis. The eccentricity is a measure of the elongation of the ellipse.

Kepler's second law, the law of areas, says that an orbiting body such as a planet sweeps out equal areas in equal amounts of time. This means that the planet's speed will be greater when it is closer to the Sun.

The angular momentum of an orbiting satellite is constant because there is no net torque on it. This fact can be used to derive Kepler's second law.

Kepler's third law, the law of periods, states that the square of the period of an orbit is proportional to the cube of the semimajor axis a, which is equal to one half the width of the orbit at its widest. For a circular orbit, a equals the radius of the orbit.

The orbital energy of a satellite is the sum of its gravitational potential energy (which is negative) and its kinetic energy. The total energy is constant, though the PE and KE change continuously if the satellite moves in an elliptical orbit.

The escape speed is the minimum speed necessary to escape a planet's gravitational attraction. It depends on the mass and radius of the planet, but not on the mass of the escaping object.

Equations

Newton's law of gravitation

$$F = \frac{GMm}{r^2}$$

Gravitational acceleration

$$g = \frac{GM}{r^2}$$

Circular orbit

$$v = \sqrt{\frac{GM}{r}}$$

Semimajor axis

$$a = \frac{r_{\text{min}} + r_{\text{max}}}{2}$$

Eccentricity

$$e = \frac{d}{a}$$

Kepler's second law

$$A = \frac{L}{2m}\Delta t$$

Kepler's third law

$$T^2 = \left(\frac{4\pi^2}{GM}\right)a^3$$

Energy in circular orbits

$$KE = \frac{GMm}{2r}$$

$$PE = -\frac{GMm}{r}$$

$$E_{\text{tot}} = -\frac{GMm}{2r}$$

Energy in elliptical orbits

$$E_{\text{tot}} = -\frac{GMm}{2a}$$

Escape speed

$$v = \sqrt{\frac{2GM}{r}}$$

Chapter 13 Problems

Chapter Assumptions

Unless stated otherwise, use the following values.

r_{Earth} = 6.38×10^6 m

M_{Earth} = 5.97×10^{24} kg

Conceptual Problems

C.2 Compare planets farther from the Sun to those nearer the Sun. (a) Do the farther planets have greater or lesser orbital speed than the nearer ones? (b) How does the angular speed of the farther planets compare to that of the nearer ones?

(a) i. Outer planets have greater orbital speed.
ii. Outer planets have smaller orbital speed.
iii. The orbital speed is the same.

(b) i. Outer planets have greater angular speed.
ii. Outer planets have smaller angular speed.
iii. The angular orbital speed is the same.

C.3 How much work is done on a satellite as it moves in a circular orbit around the Earth?

C.5 According to Kepler's third law, the ratio T^2/a^3 should be the same for all objects orbiting the Sun, since the factor $4\pi^2/GM$ is the same. When this ratio is measured however, it is found to vary slightly. For instance, Jupiter's ratio is higher than Earth's by about 1%. What are the two main assumptions behind Kepler's third law that are not 100% valid in a real planetary system?

C.6 From your intergalactic survey base, you observe a moon in a circular orbit about a faraway planet. You know the distance to the planet/moon system, and you determine the maximum angle of separation between the two and the period of the moon's orbit. Assuming that the moon is much less massive than the planet, explain how you can determine the mass of the planet.

C.7 In a previous chapter, we used the equation $PE = mgh$ to represent the gravitational potential energy of an object near the Earth. In this chapter, we use the equation $PE = -GMm/r$. Explain the reasons for the differences between these two equations: Why is one expression negative and the other positive? Is r equal to h?

C.8 Match each of the following properties of a body in circular orbit to its dependence upon the radius of the orbit: (a) velocity, (b) kinetic energy, (c) period, and (d) angular momentum.

(a) ○ r^{-1} ○ r^{-2} ○ $r^{1/2}$ ○ $r^{-1/2}$ ○ $r^{3/2}$
(b) ○ r^{-1} ○ r^{-2} ○ $r^{1/2}$ ○ $r^{-1/2}$ ○ $r^{3/2}$
(c) ○ r^{-1} ○ r^{-2} ○ $r^{1/2}$ ○ $r^{-1/2}$ ○ $r^{3/2}$
(d) ○ r^{-1} ○ r^{-2} ○ $r^{1/2}$ ○ $r^{-1/2}$ ○ $r^{3/2}$

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the first interactive problem in this section to answer the following questions. (a) Given two unchanging masses, does the force between two masses increase, decrease, or stay the same as the distance between the masses increases? (b) Given a fixed distance, does the force between two masses increase, decrease or stay the same as the masses increase?

(a) i. Increase
ii. Decrease
iii. Stay the same

(b) i. Increase
ii. Decrease
iii. Stay the same

Section 1 - Newton's law of gravitation

1.1 The Hubble Space Telescope orbits the Earth at an approximate altitude of 612 km. Its mass is 11,100 kg and the mass of the Earth is 5.97×10^{24} kg. The Earth's average radius is 6.38×10^{6} m. What is the magnitude of the gravitational force that the Earth exerts on the Hubble?

__________ N

1.2 A neutron star and a black hole are 3.34e+12 m from each other at a certain point in their orbit. The neutron star has a mass of 2.78×10^{30} kg and the black hole has a mass of 9.94×10^{30} kg. What is the magnitude of the gravitational attraction between the two?

__________ N

1.3 An asteroid orbiting the Sun has a mass of 4.00×10^{16} kg. At a particular instant, it experiences a gravitational force of 3.14×10^{13} N from the Sun. The mass of the Sun is 1.99×10^{30} kg. How far is the asteroid from the Sun?

__________ m

1.4 The gravitational pull of the Moon is partially responsible for the tides of the sea. The Moon pulls on you, too, so if you are on a diet it is better to weigh yourself when this heavenly body is directly overhead! If you have a mass of 85.0 kg, how much less do you weigh if you factor in the force exerted by the Moon when it is directly overhead (compared to when it is just rising or setting)? Use the values 7.35×10^{22} kg for the mass of the moon, and 3.76×10^{8} m for its distance above the surface of the Earth. (For comparison, the difference in your weight would be about the weight of a small candy wrapper. And speaking of candy...)

__________ N

1.5 Three 8.00 kg spheres are located on three corners of a square. Mass A is at (0, 1.7) meters, mass B is at (1.7, 1.7) meters, and mass C is at (1.7, 0) meters. Calculate the net gravitational force on A due to the other two spheres. Give the components of the force.

(__________ , __________) N

Section 2 - G and g

2.1 The top of Mt. Everest is 8850 m above sea level. Assume that sea level is at the average Earth radius of 6.38×10^{6} m. What is the magnitude of the gravitational acceleration at the top of Mt. Everest? The mass of the Earth is 5.97×10^{24} kg.

__________ m/s^2

2.2 Geosynchronous satellites orbit the Earth at an altitude of about 3.58×10^{7} m. Given that the Earth's radius is 6.38×10^{6} m and its mass is 5.97×10^{24} kg, what is the magnitude of the gravitational acceleration at the altitude of one of these satellites?

__________ m/s^2

2.3 Jupiter's mass is 1.90×10^{27} kg. Find the acceleration due to gravity at the surface of Jupiter, a distance of 7.15×10^{7} m from its center.

__________ m/s^2

2.4 A planetoid has a mass of 2.83e+21 kg and a radius of 7.00×10^{5} m. Find the magnitude of the gravitational acceleration at the planetoid's surface.

__________ m/s^2

2.5 What is the magnitude of the difference in the acceleration of gravity between the surface of the water in a swimming pool at sea level and the surface of an Olympic diving platform, 10.0 meters above? Report the answer to three significant figures.

__________ m/s^2

Section 9 - Interactive problem: Newton's cannon

9.1 Use the simulation in the interactive problem in this section to determine the initial speed required to put the cannonball into circular orbit.

__________ m/s

Section 10 - Circular orbits

10.1 The International Space Station orbits the Earth at an average altitude of 362 km. Assume that its orbit is circular, and calculate its orbital speed. The Earth's mass is 5.97×10^{24} kg and its radius is 6.38×10^{6} m.

__________ m/s

10.2 An asteroid orbits the Sun at a constant distance of 4.44e+11 meters. The Sun's mass is 1.99×10^{30} kg. What is the orbital speed of the asteroid?

__________ m/s

10.3 The Moon's orbit is roughly circular with an orbital radius of 3.84×10^{8} m. The Moon's mass is 7.35×10^{22} kg and the Earth's mass is 5.97×10^{24} kg. Calculate the Moon's orbital speed.

__________ m/s

10.4 The orbital speed of the moon Ganymede around Jupiter is 1.09×10^{4} m/s. What is its orbital radius? Assume the orbit is circular. Jupiter's mass is 1.90×10^{27} kg.

__________ m

Section 12 - Interactive problem: intercept the orbiting satellite

12.1 Use the information given in the interactive problem in this section to calculate the speed required to enter the circular orbit and destroy the rogue satellite. Test your answer using the simulation.

__________ m/s

Section 13 - Interactive problem: dock with an orbiting space station

13.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) What happens to the speed of the rocket immediately after firing its rear (Forward thrust) engine? (b) What happens to its speed after a few moments?

(a) i. It increases
ii. It decreases
iii. It stays the same

(b) i. It increases
ii. It decreases
iii. It stays the same

Section 15 - More on ellipses and orbits

15.1 A comet's orbit has a perihelion distance of 0.350 AU and an aphelion distance of 45.0 AU. What is the semimajor axis of the comet's orbit around the sun?

__________ AU

15.2 The semimajor axis of a comet's orbit is 21.0 AU and the distance between the Sun and the center of the comet's elliptical orbit is 20.1 AU. What is the eccentricity of the orbit?

____

15.3 When a planet orbits a star other than the Sun, we use the general terms periapsis and apoapsis, rather than perihelion and aphelion. The orbit of a planet has a periapsis distance of 0.950 AU and an apoapsis distance of 1.05 AU. (a) What is the semimajor axis of the planet's orbit? (b) What is the eccentricity of the orbit?

(a) __________ AU
(b) __________

15.4 The text states that the eccentricity of an elliptical orbit is equal to the distance from the ellipse's center to a focus, divided by the semimajor axis. Show that eccentricity is also equal to the positive difference in the perihelion and aphelion distances, divided by their sum.

15.5 An extrasolar planet's orbit has a semimajor axis of 23.1 AU. The eccentricity of the orbit is 0.010. What is the periapsis distance (the planet's minimum distance from the star it is orbiting)?

__________ AU

15.6 The Trans-Neptunian object Sedna was discovered in 2003. By mid-2004, Sedna's orbit was estimated to have a semimajor axis of 480 AU and an eccentricity of 0.84. (a) What is the perihelion distance of Sedna's orbit? (b) What is the aphelion distance?

(a) __________ AU
(b) __________ AU

Section 18 - Kepler's third law

18.1 Jupiter's semimajor axis is 7.78×10^{11} m. The mass of the Sun is 1.99×10^{30} kg. (a) What is the period of Jupiter's orbit in seconds? (b) What is the period in Earth years? Assume that one Earth year is exactly 365 days, with 24 hours in each day.

(a) __________ s
(b) __________ years

18.2 Mars orbits the Sun in about 5.94×10^{7} seconds (1.88 Earth years). (a) What is its semimajor axis in meters? (The mass of the Sun is 1.99×10^{30} kg.) (b) What is Mars' semimajor axis in AU? 1 AU = 1.50×10^{11} m.

(a) __________ m
(b) __________ AU

18.3 A planet orbits a star with mass 2.61e+30 kg. The semimajor axis of the planet's orbit is 2.94×10^{12} m. What is the period of the planet's orbit in seconds?

__________ s

18.4 An extrasolar planet has a small moon, which orbits the planet in 336 hours. The semimajor axis of the moon's orbit is 1.94e+9 m. What is the mass of the planet?

__________ kg

18.5 Jupiter's moon Callisto orbits the planet at a distance of 1.88×10^{9} m from the center of the planet. Jupiter's mass is 1.90×10^{27} kg. What is the period of Callisto's orbit, in hours?

__________ hours

18.6 Write Kepler's third law for planets in the solar system, with *T* measured in years and *a* in astronomical units (AU), and explain how you arrived at the equation.

18.7 The Trans-Neptunian object Sedna has an extremely large semimajor axis. In the year 2004, it was estimated to be 480 AU. What is the period of Sedna's orbit, measured in Earth years?

__________ years

Section 20 - Interactive problem: geosynchronous satellite

20.1 Use the simulation in the interactive problem in this section to calculate the (a) radius and (b) launch speed required to achieve geosynchronous orbit around the Earth.

(a) __________ km
(b) __________ m/s

Section 21 - Orbits and energy

21.1 The Hubble Space Telescope orbits the Earth at an altitude of approximately 612 km. Its mass is 11,100 kg and the mass of the Earth is 5.97×10^{24} kg. The Earth's radius is 6.38×10^{6} m. Assume the Hubble's orbit is circular. (a) What is the gravitational potential energy of the Earth-Hubble system? (Assume that it is zero when their separation is infinite.) (b) What is the Hubble's *KE*? (c) What is the Hubble's total energy?

(a) __________ J
(b) __________ J
(c) __________ J

21.2 What is the least amount of energy it takes to send a spacecraft of mass 3.50×10^{4} kg from Earth's orbit to that of Mars? (Neglect the gravitational influence of the planets themselves.) Assume that both planetary orbits are circular, the radius of Earth's orbit is 1.50×10^{11} meters, and that of Mars' orbit is 2.28×10^{11} meters. The Sun's mass is 1.99×10^{30} kg.

__________ J

21.3 How much work is required to send a spacecraft from Earth's orbit to Saturn's orbit around the sun? The semimajor axis of Earth's orbit is 1.50×10^{11} meters and that of Saturn's orbit is 1.43×10^{12} meters. The spacecraft has a mass of 3.71e+4 kg and the Sun's mass is 1.99×10^{30} kg.

__________ J

21.4 You wish to boost a 9,550 kg Earth satellite from a circular orbit with an altitude of 359 km to a much higher circular orbit with an altitude of 35,800 km. What is the difference in energy between the two orbits, that is, how much energy will it take to accomplish the orbit change? Earth's radius is 6.38×10^{6} m and its mass is 5.97×10^{24} kg.

__________ J

21.5 A satellite is put in a circular orbit 485 km above the surface of the Earth. After some time, friction with the Earth's atmosphere causes the satellite to fall to the Earth's surface. The 375 kg satellite hits the Pacific Ocean with a speed of 2,500 m/s. What is the change in the satellite's mechanical energy? (Watch the sign of your answer.) In this situation, mechanical energy is transformed into heat and sound. The Earth's mass is 5.97×10^{24} kg, and its radius is 6.38×10^{6} m.

__________ J

21.6 You launch an engine-less space capsule from the surface of the Earth and it travels into space until it experiences essentially zero gravitational force from the Earth. The initial speed of the capsule is 18,500 m/s. What is its final speed? Assume no significant gravitational influence from other solar system bodies. The Earth's mass is 5.97×10^{24} kg, and its radius is 6.38×10^{6} m.

__________ m/s

Section 25 - Interactive problem: a rocket mission to Mars

25.1 Use the information given in the interactive problem in this section to calculate the energy required to propel a rocket from Earth's circular orbit into an elliptical transfer orbit that will reach Mars. Test your answer using the simulation.

__________ J

Section 26 - Escape speed

26.1 Calculate the escape speed from the surface of Venus, whose radius is 6.05×10^{6} m and mass is 4.87×10^{24} kg. Neglect the influence of the Sun's gravity.

__________ m/s

26.2 The escape speed of a particular planet with a radius of 7.60e+6 m is 14,500 m/s. What is the mass of the planet?

__________ kg

26.3 A planet has a mass of 5.69×10^{26} kg. The planet is not a perfect sphere, instead it is somewhat flattened. At the equator, its radius is 6.03×10^{7} m, and at the poles, the radius is 5.38×10^{7} m. (a) What is the escape speed from the surface of the planet at the equator? (b) What is the escape speed at the poles?

(a) __________ m/s
(b) __________ m/s

26.4 The average density of Neptune is 1.67×10^{3} kg/m^3. What is the escape speed at the surface? Neptune's radius is 2.43×10^{7} m.

__________ m/s

26.5 The Schwarzschild radius of a black hole can loosely be defined as the radius at which the escape speed equals the speed of light. Anything closer to the black hole than this radius can never escape, because nothing can travel faster than light in a vacuum. Not even light itself can escape, hence the name "black hole". (a) Find the Schwarzschild radius of a black hole with a mass equal to five times that of the Sun. This is a typical value for a "stellar-mass" black hole. The Sun's mass is 1.99×10^{30} kg. (b) Find the Schwarzschild radius of a black hole with a mass of one billion Suns. This is the type of black hole found at the centers of the largest galaxies.

(a) __________ m
(b) __________ m

Additional Problems

A.1 A satellite in geosynchronous orbit does not change its position in the sky as seen from Earth. This means it orbits the Earth once each day, moving in the same direction as the Earth's rotation. (a) What is the altitude of a satellite in geosynchronous orbit? (The radius of the Earth is 6.38×10^{6} m and its mass is 5.97×10^{24} kg.) (b) What is its speed?

(a) __________ m
(b) __________ m/s

A.2 The radius of Mars is 0.533 times that of the Earth and the gravitational acceleration on Mars' surface is 0.378 times that on the Earth's surface. What is the ratio of Mars' average density to the Earth's average density?

____

A.3 A certain binary star system consists of two identical stars in circular orbits about a common center of mass halfway between them. Their orbital speed is 185,000 m/s and each one orbits the center of mass in exactly 19 days. What is the mass of each star, in units of solar masses? The mass of the Sun is 1.99×10^{30} kg. Hint: Equate the gravitational force to the centripetal force causing circular motion.

__________ solar masses

A.4 The Chandra X-ray Observatory is a telescope in an elliptical Earth orbit. The elongated orbit takes the telescope out of the Van Allen radiation belts surrounding the Earth for 85% of its orbit. (The charged particles in the belts would otherwise interfere with Chandra's X-ray observations.) Chandra's orbit takes it to a maximum altitude of 133,000 km and a minimum altitude of 16,000 km. (a) What is the semimajor axis of Chandra's orbit? (b) What is the eccentricity of the orbit? (c) What is the total energy associated with this orbit? The radius of the Earth is 6380 km, the mass of the Earth is 5.97×10^{24} kg and Chandra's mass is 4,800 kg.

(a) __________ km
(b) __________
(c) __________ J

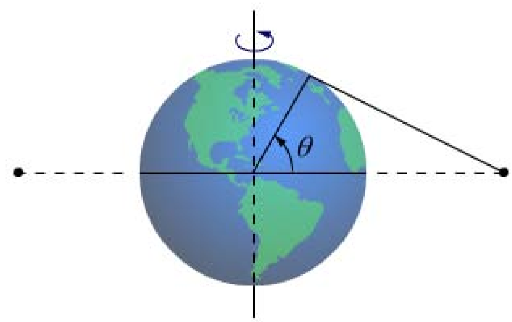
A.5 (a) Calculate the escape speed of a rocket from Mercury's gravity. Mercury's radius is 2.44×10^6 m and its mass is 3.30×10^{23} kg. (b) Suppose a rocket takes off from the side of Mercury directly opposite the Sun. Calculate the escape speed, accounting for the gravitational influence of both Mercury and the Sun. The Sun's mass is 1.99×10^{30} kg and the center-to-center distance from Mercury to the Sun is 5.79×10^{10} m. Hint: You cannot start with the escape speed equation given in the book. You must start with the conservation of energy to derive an expression for the escape speed for this case. You should find that the answers in parts a and b are significantly different.

(a) __________ m/s
(b) __________ m/s

A.6 In June 2002, scientists at Caltech discovered a new orbiting body in the solar system, half the diameter of Pluto. Quaoar (KWAH-o-ar) takes 288 years to complete one orbit around the sun, and its orbit is remarkably circular. Find the distance from Quaoar to the sun (the sun's mass is 1.99×10^{30} kg). Assume a year has 365 days.

__________ m

A.7 Satellites in geosynchronous orbit (which remain above the same point on Earth) must be positioned directly above the equator, which is zero degrees latitude. These satellites are not visible from locations near the poles. What is the most extreme latitude from which such a satellite can still be seen? Refer to the illustration, where the angle θ is equal to the latitude of the point where the radial line intersects the Earth's surface. The radius of the Earth is 6.38×10^6 m and its mass is 5.97×10^{24} kg.



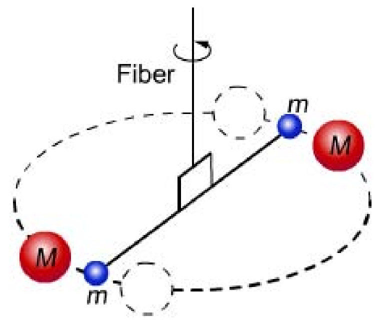
__________ degrees

A.8 A geosynchronous satellite is stationary over a point on the equator (zero degrees latitude) at the same longitude as Seattle, Washington. Seattle's latitude is 47.6°. If you want to communicate with the satellite, at what angle above the horizon must you point your communication device? The Earth's radius is 6.38×10^6 m. The orbital radius of a geosynchronous satellite was found in a previous problem (you can also calculate it using the mass of the Earth: 5.97×10^{24} kg). Hint: You will also need the law of sines and the law of cosines.

__________ degrees

A.9 Kepler's third law as stated in the text assumes that the mass of the central body is much greater than that of the satellite. A more general version of the law is $T^2/a^3 = 4\pi^2/(G(M + m))$, which includes a term for the mass m of the satellite. Derive this expression assuming that both bodies are in circular orbit about a common center of mass. The term a represents the constant separation of the two bodies, and reduces to the semimajor axis in the case when one body is much larger than the other. Hint: Equate the gravitational force to the centripetal force causing circular motion.

A.10 In the late 18th century, Henry Cavendish used an apparatus like the one pictured, called a torsion balance, to measure the gravitational constant *G*. The small masses *m* are attached to a light rod of length *L* which is suspended by a fiber. The large masses *M*, when placed at the positions shown, attract the smaller masses. This exerts a torque on the rod and the fiber twists. As the fiber twists, it exerts a restoring torque on the rod, whose magnitude equals the torsional constant κ multiplied by the magnitude of the fiber's angular displacement from equilibrium. When the two torque balance, the system is in equilibrium. Derive an equation for *G* in terms of θ, *M*, *m*, *L*, *R* (the distance between the centers of each *M* and *m* pair in equilibrium) and κ (the torsional constant of the fiber). Assume that each large ball only applies a gravitational force on the small ball that is closest to it. If you have to perform the experiment, it is helpful to note that when the large masses are moved to the alternate, outlined positions, the rod experiences a torque in the opposite direction and moves through an angle 2θ until it is again in equilibrium.

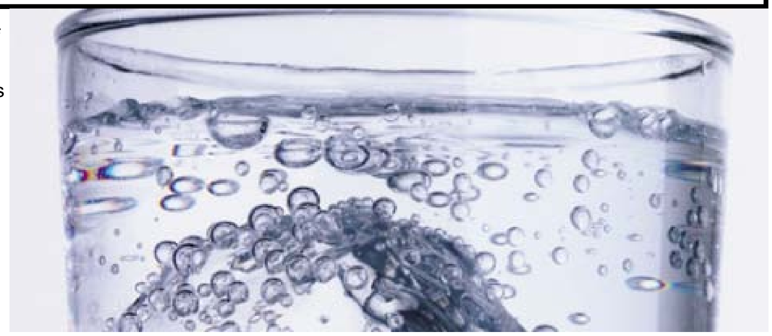


- $G = \kappa\theta RL/Mm$
- $G = \kappa\theta R^2/MmL$
- $G = Mm\theta/\kappa R^2L$
- $G = \kappa\pi\theta L^2/MmR$

A.11 You wish to launch a rocket straight up from the Earth. By taking advantage of the tangential velocity due to the Earth's rotation, you can reduce the vertical velocity required to achieve escape speed. The greatest advantage can be achieved at the equator. (a) What is the tangential velocity at the equator due to Earth's rotation? (The radius of the Earth is 6.378×10^6 m.) (b) What is the magnitude of the vertical velocity required to achieve escape speed at the equator? Hint: Set the escape speed equal to the magnitude of the vector sum of the tangential and vertical velocities. (The mass of the Earth is 5.974×10^{24} kg. The gravitational constant G equals 6.674×10^{-11} N·m²/kg².) (c) How much more vertical speed is required if the rocket is launched straight up from the North or South Pole? (Assume that the Earth is a perfect sphere.)

(a) __________ m/s
(b) __________ m/s
(c) __________ m/s

14.0 - Introduction



The glass and the ice cube in this photo are solid objects. The water is a liquid. The bubbles and the surrounding air are gasses.

The study of physics typically begins with the study of solid objects: You learn how to determine the velocity of a car as it accelerates down a street, what happens when two pool balls collide, and so on.

This chapter introduces the study of fluids. Liquids and gases are both fluids. Fluids change shape much more readily than solids. Pour soda from a can into a glass and the liquid will change shape to conform to the shape of the glass. Push on a balloon full of air or water and you can easily change the shape of the balloon and the fluid it contains. In contrast to liquids, gases expand to fill all the space available to them. One reason astronauts wear spacesuits is to keep their air near them, and not let it expand limitlessly into the near vacuum of space.

This chapter focuses on the characteristics exhibited by fluids when their temperature and density remain nearly constant. It covers topics such as the method of calculating how much pressure water will exert on a submerged submarine, and why a boat floats. Some of the topics apply to liquids alone, while others apply to both liquids and gases.

14.1 - Fluid

Fluid: A substance that can flow and conform to the shape of a container. Liquids and gases are fluids.

A fluid alters its shape to conform to the shape of the container that surrounds and holds it. The molecules of a fluid can "flow" because they are not fixed into position as they would be in a solid. Liquids and gases are fluids, and they are two of the common forms of matter, with solids being the third. There are other forms of matter as well, such as plasma (created in fusion reactors) and degenerate matter (found in neutron stars).

A substance can exist as a solid, a liquid or a gas depending on the surrounding physical conditions. Factors such as temperature and pressure determine its state. For the purposes of a physics textbook, we need to be specific about what state of matter we are discussing at any given time. However, whether a substance is considered a solid or a fluid may depend on factors such as the time scale under consideration. For instance, the ice in a glacier can seem quite solid, but glaciers do flow slowly over time, so treating glacier ice as a fluid is useful to geologists.

concept 1

Fluids
Can flow
Rate of flow varies

concept 2

Fluids
Conform to container

14.2 - Density

Density: Mass per unit volume.

Density − to be precise, *mass density* − equals mass divided by volume. The Greek letter ρ (rho, pronounced "roe") represents density. The SI unit for density is the kilogram per cubic meter. The gram per cubic centimeter is also a common unit, useful in part because the density of water is close to one gram per cubic centimeter.

Liquids and solids retain a fairly constant density. It requires a great deal of force to compress water or a piece of clay into a more compact form. A given mass of liquid will change shape in order to conform to the shape of the container you pour it into, but its volume will remain nearly constant over a great variety of conditions.

In contrast, the volume of a sample of gas changes readily, which means its density changes easily, as well. For example, when you pump a stroke of air into a bicycle tire, the volume of air in the pump cylinder is reduced to "squeeze" it into the tire, increasing the air's density.

Much larger changes can be accomplished using larger increases in pressure. Machinery compresses the air fed into a scuba diving tank, causing its contents to be at a density on the order of 200 times greater than the density of the atmosphere you breathe. Before a diver breathes this air, its density (and pressure) are reduced. It would be lethal at the pressure maintained in the tank.

The density of an object can vary at different points based on its composition. A precise way to state the definition of density is $\Delta m/\Delta V$: The mass of a small volume of material is measured to establish its local density. However, unless otherwise stated, we will assume that the substances we deal with have *uniform density*, the same density at all points. This means that the density can be established by dividing the total mass by the total volume.

The table on the right shows the densities of various materials. Their densities are given at $0°\text{C}$ and one standard atmosphere of pressure (the density of the air around you can vary, depending on atmospheric conditions). Exceptions to this are the super-dense neutron star, which has no temperature in the ordinary sense, and liquid water, whose density is given at $4°\text{C}$, the temperature at which it is the greatest. We also supply some common equivalents for water, for instance relating a liter of water to its mass.

The concept of *specific gravity* provides a useful tool for understanding and comparing various materials' densities. Specific gravity divides the density of one material by that of a reference material, usually water at $4°\text{C}$. For instance, if a material has a specific gravity of two, it is twice as dense as water.

concept 1

Density
Ratio of mass to volume

equation 1

Density

$$\rho = m/V$$

ρ = density
m = mass
V = volume
Units: kg/m^3

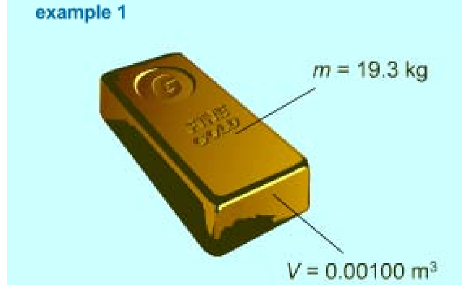
equation 2

Substance*	Density in kg/m^3
Neutron star	5×10^{17}
Mercury	13,595
Lead	11,300
Aluminum	2700
Water (4° C)	1000
Ice (0° C)	917
Hydraulic oil	890
Air	1.28
Hydrogen gas	0.082

* most at 0° C. Gases at 1.013×10^5 Pa

Density of various substances
For water at 4° C:
· 1 liter has mass of 1 kg
· 1 cm^3 has mass of 1 g

example 1



What is the density of the gold brick?

$\rho = m/V$

ρ = 19.3 kg/0.00100 m^3

ρ = 19,300 kg/m^3

14.3 - Pressure

Pressure: Final exams, SATs, free throws in the last 30 seconds of a tight game, and driver's license tests.

Pressure: Force divided by the surface area over which the force acts.

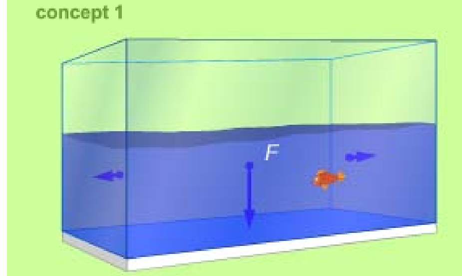
Big feet distribute this lynx's weight over a larger surface area, lowering the pressure they exert on the snow.

The first definition of pressure above speaks for itself, but the second deserves further explanation.

You experience pressure when you swim. If you dive deep under the water, you can feel the water pushing against you with more pressure, more force per square meter of your body.

For a surface immersed in a fluid, the amount of pressure at a given location in the fluid is the same for any orientation of the surface. If you place your hand thirty centimeters underwater, the pressure on your palm is the same no matter how you rotate it. In the aquarium illustrated on the right, the water exerts a force on the bottom of the tank, but it exerts force − and pressure − on the sides as well.

concept 1



Pressure
Force divided by surface area

Pressure equals the amount of force divided by the surface area to which it is applied. As the photograph above shows, some animals, such as the lynx, are able to travel easily across the surface of snow because their large paws spread the force of their weight over a large area. This reduces the pressure they exert on the snow, enabling them to walk on its surface. Animals of similar weight, but with smaller paws, sink into the snow.

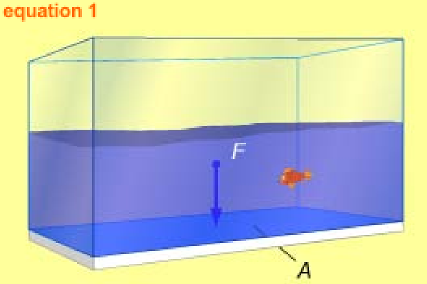
People who need to travel across the snow may likewise use snowshoes or skis to increase surface area and reduce the pressure they exert. In contrast, spike-heeled shoes concentrate almost all the weight of their wearers over a very small surface area, and can exert a pressure large enough to damage wood or vinyl floors.

The water in the aquarium exerts force on the walls of the tank as well as its bottom. This is shown in Concept 1. Why does the water exert a force on the sides of the aquarium? Consider squishing down on a water balloon: The balloon bulges out on its sides. The additional force you exert on the top is translated into a force on the sides. To return to the aquarium, the downward weight of the water results in a force on the walls as well as the bottom.

The formula in Equation 1 shows how to calculate pressure. It equals the magnitude of the force divided by the area and is a scalar quantity. The SI unit for pressure is the newton per square meter, called a *pascal* (Pa). One pascal is a very small amount of pressure. The Earth's atmosphere exerts about 100,000 pascals of air pressure at the planet's surface. A *bar* of pressure equals 100,000 (10^5) pascals and is another commonly used unit. Bars are informally called "atmospheres" (atm) of pressure. You may have heard references to "millibars" in weather reports.

In the British system, force is measured in pounds. Pounds per square inch, *psi*, is a common measure of force. At the Earth's surface, typical atmospheric pressure is 14.7 psi. Automobile tires are normally inflated to a pressure of about 24 psi, while road bicycle tires are inflated to pressures as high as 120 psi. The tire readings reflect the amount of pressure inside the tires **over** atmospheric pressure. The total (absolute) pressure inside a tire is the sum of atmospheric pressure and the gauge reading (a total of about 135 psi for the bicycle tire).

equation 1



Pressure

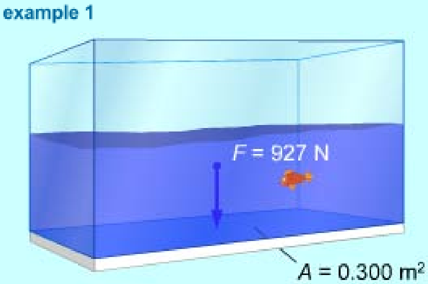
$P = F/A$

P = pressure
F = force
A = surface area
Units: pascal (Pa), newton/meter^2

equation 2

Atmospheric pressure at sea level
P_{atm} = 101,300 Pa ≈ 1 bar

example 1



What is the pressure inside the bottom of the aquarium caused solely by the weight of the water?
$P = F/A$
P = 927 N / 0.300 m^2
P = 3090 Pa

14.4 - Pressure and fluids

Navy personnel performing under pressure.

As a submarine dives deeper and deeper into the sea, it encounters increasing pressure. The nightmare of any submariner is that his craft goes so deep it collapses under the tremendous pressure of the water. (Rent the movie *U-571* if you want to watch a Hollywood thriller that deals with pressure at great depths.)

Physicists prefer a little less drama and a little more measurement in their dealings with pressure. They have developed an equation to describe the pressure of a fluid as a function of the fluid's depth.

Why does water pressure increase as a submarine descends? The pressure increases because the amount of water on top of the submarine increases as the vessel does down. The weight of the water exerts a force over the entire surface area of the submarine's hull. The water pressure does not depend on the orientation of the surface. It exists all over the craft: on its top, bottom and sides.

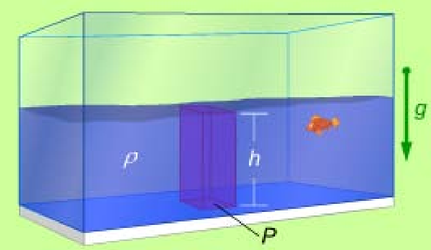
The first equation on the right shows how to calculate the pressure at a point underneath the surface of a fluid. The pressure equals the product of the fluid's density, the constant acceleration of gravity, and the height of the column of fluid above the point. "Height" means the distance from the point to the surface of the fluid. The equation states that pressure increases linearly with depth. The water pressure at 200 meters down is twice as great as it is at 100 meters. This equation applies when the fluid is static.

The fluid pressure varies solely as a function of the density ρ and the depth in the fluid, not with the shape of the container holding the fluid. If you fill a swimming pool and a Coke bottle with water, the water pressure at 0.1 meters below the surface will be the same in both containers. The pressure will also be the same at the bottom of the bottle and on the wall of the pool at the same depth. Whether the surface area is horizontal or vertical, the pressure is the same.

You can add pressures. The **total** pressure exerted on the exterior of the hull of the submarine equals the sum of *atmospheric pressure* (the pressure exerted by the Earth's air above it) and the pressure of the water above it. The two pressures must be calculated separately and added as shown in Equation 2. Since the densities of water and air differ, the pressure of the combined column of fluids above the submarine cannot be calculated as a single product ρgh.

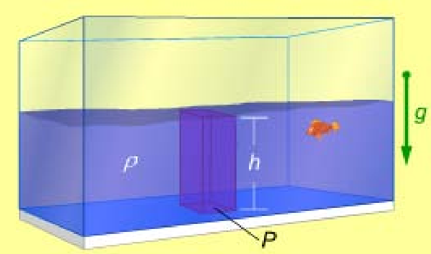
We have implicitly described two types of pressure: absolute and gauge pressure. The total pressure is called the *absolute pressure*. It is the sum of the atmospheric pressure and the pressure of the fluid in question, in this case water. The photograph below shows how a Styrofoam® cup (as shown on the right) was crushed (as shown on the left) by an absolute pressure of 3288 psi when it was submerged to a depth of more than two kilometers below the ocean's surface.

concept 1



Pressure due to a fluid
Function of:
· density of fluid
· acceleration of gravity
· height of column

equation 1



Pressure of a fluid (liquid)

$$P = \rho gh$$

P = pressure
ρ = density of liquid
g = acceleration of gravity
h = height of liquid column

The term *gauge pressure* describes the pressure caused solely by the water (or other fluid), ignoring the atmospheric pressure. The gauge pressure equals ρgh, where ρ, g and h are measured for the fluid alone. To state the same concept in another way: Gauge pressure equals the absolute pressure minus the atmospheric pressure.

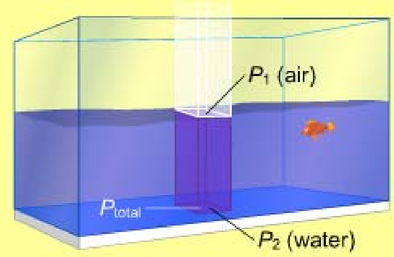
Pressures can oppose one other, when they act on opposite sides of the same surface. For example, the "atmospheric" pressure inside an airplane cabin is allowed to decrease as the plane climbs. The pressure inside your ear can be higher than the pressure of the cabin, since it may remain at the higher pressure of the atmosphere at the Earth's surface. In this case, the pressure in parts of your ears is greater than the pressure outside them, producing a net outward pressure (and force) on your eardrums. The result is that your ear begins to ache.

You can reduce the pressure inside your ears by chewing gum or yawning to "pop" them. This stretches and opens the Eustachian tubes, passages between your ears and throat. Air flows out of your ear into the cabin, balancing the pressures, and your earache disappears.

If you look at the large value calculated in the example problem for the absolute pressure on the inner surface of the aquarium's bottom, you might wonder why the tank does not burst. Remember that atmospheric pressure is pushing inward on its exterior surfaces as well. The net pressure on the bottom plate is the gauge pressure due solely to the water, which is not large enough to cause the tank to burst.

The density of a liquid does not vary significantly with depth, so using an average density figure in the formula $\rho g h$ provides a good approximation of the pressure at any depth. The density of gases can vary greatly, so the formula is not as applicable to them, especially if h is large. For instance, the density of the Earth's atmosphere is about 1.28 kg/m^3 at sea level (and $0°\mathrm{C}$), but only 0.38 kg/m^3 at an altitude of 10,600 m (35,000 ft) above sea level (at $-20°\mathrm{C}$). The lower layers of the atmosphere are significantly compressed by the weight of the air above them. Mountain climbers note this difference, remarking that the air is "thinner" at higher altitudes.

equation 2

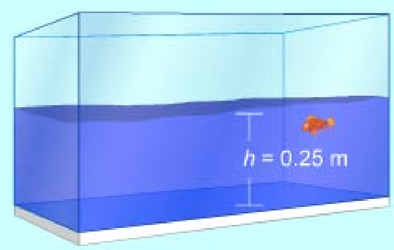


Pressure can be summed

$$P_{total} = \Sigma P = P_1 + P_2 + \ldots + P_n$$

Total pressure = sum of pressures

example 1



What is the gauge pressure inside the bottom of the tank due to the weight of the water alone? What is the absolute pressure, including atmospheric pressure, pressing down inside the bottom?

$P_{H2O} = \rho g h$

$P_{H2O} = (1000\ \frac{\mathrm{kg}}{\mathrm{m}^3})(9.80\ \frac{\mathrm{m}}{\mathrm{s}^2})(0.25\ \mathrm{m})$

$P_{H2O} = 2400\ \mathrm{Pa}$

$P_{abs} = 2400\ \mathrm{Pa} + 101{,}300\ \mathrm{Pa}$

$P_{abs} = 103{,}700\ \mathrm{Pa}$

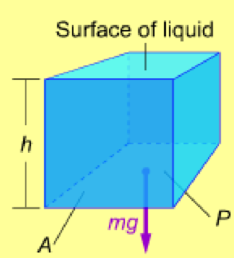
14.5 - Derivation: liquid pressure

We use the concept of forces in equilibrium to derive the equation for pressure in a liquid. The liquid is assumed to have a uniform density ρ, no matter what its depth.

In the illustration to the right we show a column of liquid, such as water. The column extends from the surface of the liquid down to the depth where the pressure is to be determined. It has height h and its bottom surface has area A. The mass of the liquid in the column is m.

Two forces act on the column. Since we are interested in the gauge pressure at depth h due to the liquid alone, we ignore the effect of atmospheric pressure on its top surface. The pressure P of the liquid below the column presses up on its bottom surface with a force PA. Gravity acting on its mass pulls it down; the magnitude of this force is the column's weight, mg. We assume that the liquid is static, so we know that the column is not accelerating in any direction.

equation 1



Liquid pressure

$$P = \rho g h$$

P = pressure at depth h

ρ = density of liquid

Variables

mass of column	m
acceleration of gravity	g
pressure on bottom surface	P
area of bottom surface	A
density of liquid	ρ
height of column	h

g = acceleration of gravity
h = depth

Strategy

1. Since the column is not accelerating, the net force acting on it must be zero. This means the forces acting on it must sum to zero.
2. Use the definitions of density, weight and volume to rearrange and simplify the equation that sums the forces to zero.

Physics principles and equations

Newton's second law

$$\Sigma F = ma$$

The definitions of pressure and weight

$$P = F/A \qquad weight = mg$$

The definition of density

$$\rho = m/V$$

The volume of the column equals the area of its base times its height:

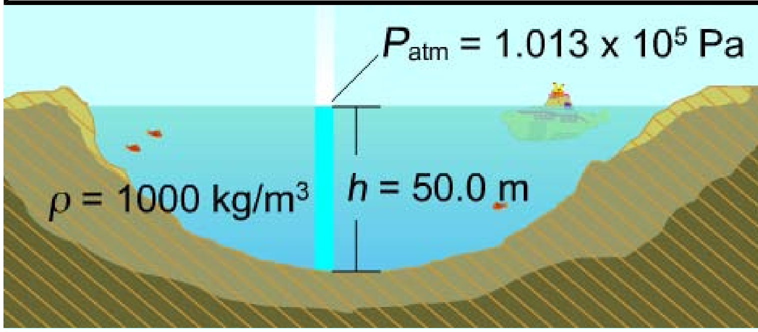
$$V = Ah$$

Step-by-step derivation

Since the column of liquid is not accelerating, Newton's second law implies that the upward and downward forces on it are equal.

Step	Reason
1. $\Sigma F = ma = 0$	Newton's 2nd law; column not accelerating
2. $PA + (-mg) = 0$	definitions of pressure and weight
3. $PA = mg$	rearrange
4. $PA = (\rho Ah)g$	definition of density
5. $P = \rho gh$	simplify

14.6 - Sample problem: pressure at the bottom of a lake



This lake is at sea level. What is the absolute pressure at the bottom of the lake?

Variables

absolute pressure	P_{abs}
atmospheric pressure	$P_{atm} = 1.013\times10^5$ Pa
gauge pressure due to water	P_{H2O}
density of water	$\rho = 1000$ kg/m^3
acceleration of gravity	$g = 9.80$ m/s^2
height of water column	$h = 50.0$ m

Strategy

1. The atmospheric pressure is given. Find the gauge pressure of the water in the lake.
2. Add the atmospheric and gauge pressures to get the absolute pressure.

Physics principles and equations

The gauge pressure of the water at the bottom of the lake is

$$P_{H2O} = \rho g h$$

The absolute pressure at the bottom of the lake is

$$P_{abs} = P_{H2O} + P_{atm}$$

Step-by-step solution

P_{atm} is known. We find P_{H2O} and add the two quantities together to find the absolute pressure P_{abs}.

Step	Reason
1. $P_{atm} = 1.013\times10^5$ Pa	standard value
2. $P_{H2O} = \rho g h$	gauge pressure equation
3. $P_{H2O} = (1000\ \frac{kg}{m^3})(9.80\ \frac{m}{s^2})(50.0\ m)$	substitute values
4. $P_{H2O} = 4.90\times10^5$ Pa	evaluate
5. $P_{abs} = P_{H2O} + P_{atm}$	absolute pressure
6. $P_{abs} = 5.91\times10^5$ Pa $P_{abs} \approx 6$ atm (bars)	substitute equations 1 and 4 into equation 5

Atmospheric pressure is not constant, even at a fixed height like sea level. In Concept 1, you see a simple *barometer*. Barometers measure atmospheric pressure. The terms “barometer” and “barometric pressure” are sometimes heard in weather reports because changes in atmospheric pressure often indicate that a change of weather is on the way.

Dial-type barometer, calibrated in millimeters of mercury and millibars.

The barometer depicted on the right consists of a vertical tube, closed at the top. This tube is partially filled with a liquid, commonly the liquid metal mercury, and is placed in a container that serves as a reservoir. The mercury in the tube stands in a column, with the space at the top of the tube occupied by a near vacuum, which exerts negligible pressure.

Increased air pressure pushing down on the surface of the reservoir causes the column of mercury to rise, until the air pressure and the pressure due to the mercury column reach equilibrium. The pressure exerted by this liquid column, the product of its density, the acceleration of gravity, and its height, equals the external air pressure.

On a typical day at sea level, air pressure causes the mercury to rise to a height of about 760 millimeters, or about 30 inches. (The pressure of a one-millimeter column of mercury is called a *torr*, after the physicist Evangelista Torricelli.) The design of this instrument should give you a sense of how strong atmospheric pressure is: It is able to force a column of mercury, which is denser than lead, to rise more than three-quarters of a meter.

In Concept 2 you see an *open-tube manometer*, a device for measuring the gauge pressure of a gas confined in a spherical vessel. The vessel that contains the gas is connected to a U-shaped tube partially filled with mercury and open to the atmosphere at its far end. This apparatus allows physicists to accurately determine the gauge pressure of the gas.

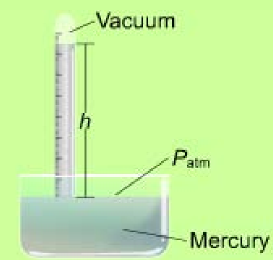
The pressure inside the spherical vessel on the left-hand side is an absolute pressure. It presses down on the surface of the left-hand column of mercury, but the higher column of mercury and the air pressure on the right side push back. When the two columns of mercury are in equilibrium, the absolute pressure of the gas equals the sum of the atmospheric pressure and the pressure exerted by the extra mercury on the right-hand side of the instrument. This equality is shown in Equation 2. In the equation, the product ρgh represents the pressure exerted by a column of mercury of height h.

The height h of the extra mercury, indicated in the diagram, equals the amount by which the mercury level on the right is higher than the mercury level on the left. It can be used with the density of mercury and the acceleration of gravity to calculate the product ρgh. This product equals the difference between the absolute pressure of the gas in the vessel and atmospheric pressure. In other words, the product ρgh gives the gauge pressure of the gas, in pascals. When the gauge pressure is expressed in “millimeters of mercury,” or torr, then its numeric value equals the numeric value of the height h, measured in millimeters.

Traditional blood pressure gauges (*sphygmomanometers*) are open-tube manometers. The gauge has an inflatable cuff that is wrapped around your upper arm. Like the sphere in the diagram, the cuff can be filled with pressurized air. This restricts the flow of blood to the lower parts of your arm. Air is then released from the cuff until the first flow of blood can be heard with a stethoscope. At this point, the gauge pressure of the blood being pumped by your heart equals the cuff’s gauge pressure. This blood pressure, the *systolic pressure*, occurs when the heart generates its maximum pressure.

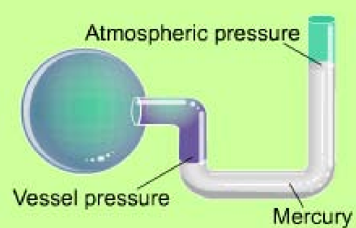
The sphygmomanometer operator (try saying that quickly!) then listens for the part of the heartbeat cycle when the pressure is the lowest, releasing pressure from the cuff until its pressure is the same as the lowest blood pressure, and the flow of blood can be heard continuously. This lower pressure is the *diastolic pressure*. A young, healthy human has a systolic blood gauge pressure of about 120 millimeters of mercury and a diastolic pressure equal to about 80 millimeters of mercury.

concept 1



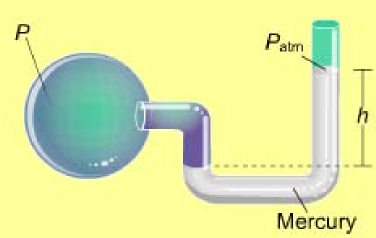
Barometer
Air pressure equals pressure of mercury
· Height of mercury reflects air pressure

concept 2



Open-tube manometer
Vessel pressure
= atm pressure + mercury pressure

equation 1



Open-tube manometer

$$P = P_{atm} + \rho gh$$

P = absolute pressure in vessel
P_{atm} = atmospheric pressure
ρgh = gauge pressure

example 1

The absolute pressure in the sphygmomanometer cuff is 1.177×10^5 Pa. What is the man's blood pressure reading, in torr?

$$P = P_{atm} + \rho g h$$

$$h = \frac{P - P_{atm}}{\rho g}$$

$$h = \frac{1.177 \times 10^5 \text{ Pa} - 1.013 \times 10^5 \text{ Pa}}{(13{,}600 \text{ kg/m}^3)(9.80 \text{ m/s}^2)}$$

$h = 0.123$ m = 123 mm

systolic blood pressure = 123 torr

14.8 - Interactive checkpoint: a 777

A Boeing 777-200 commercial jetliner has 12 main tires under the middle of the fuselage that together support 92.0% of the plane's taxi weight of 3.42×10^6 N (the nose gear supports the rest of the weight). Each tire is slightly flattened by its contact with the ground, so that the area of contact is 53.3 cm (the width of one tire) by 32.8 cm. What is the gauge pressure of the tires? Assume that each of the 12 main tires supports an equal portion of the weight.

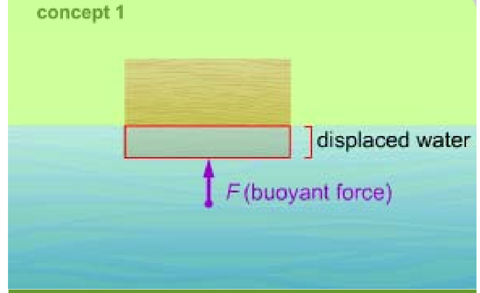
Answer:

$P =$ ______ Pa

14.9 - Archimedes' principle

Archimedes' principle: An object in a fluid experiences an upward force equal to the weight of the fluid it displaces.

concept 1



Archimedes' principle
Buoyant force:
· upward force on an object in a fluid

Archimedes (287-212 BCE) explained why objects float. His principle states that *buoyancy*, the upward force caused by the displacement of fluid, equals the weight of the volume of the fluid displaced. For instance, if a boat displaces 300 tons of water, then it experiences an upward buoyant force of 300 tons. If this buoyant force is greater than the weight of the boat, the boat floats.

Archimedes' principle can be used to explain why a small stone sinks, while a large block of Styrofoam® floats, even if the Styrofoam block is much heavier than the stone. Stone is denser than water, so the weight of the water it displaces is less than its own weight. This means that the stone's weight, directed down, is greater than the upward buoyant force on it. The net force on an underwater stone is downward.

In contrast, Styrofoam is less dense than water. A Styrofoam block displaces a weight of water equal to its own weight when it is only partially submerged. It is in a state of equilibrium since the buoyant force up equals the weight down. In short, it floats.

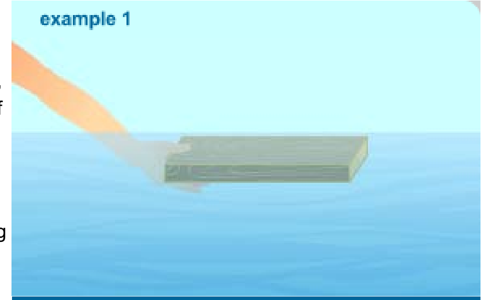
Archimedes' principle applies at any water depth. The buoyancy of a submerged submarine is the same whether it is 50 meters or 300 meters below the water's surface, since the density (and weight) of the displaced water changes only slightly with depth. If this is so, how can a submarine dive or surface? The submarine changes its weight: It either adds weight by allowing water into its ballast tanks or reduces its weight by blowing the water out with compressed air.

Fish approach the issue in a slightly different way. They change their volume by inflating or deflating an organ called a swim bladder, filled with gas released from the blood of the fish. When they increase their volume, they rise because they displace more water and experience increased upward buoyancy.

Archimedes' principle can also be used to analyze the buoyancy of human beings. People with a high percentage of body fat float more easily than do their slimmer counterparts. This is because fat is less dense than water, while muscle is denser than water. In one test for lean body mass, a person is weighed out of water, and then weighed again while submerged. The difference in the two weights equals the buoyant force, which allows a calculation of the volume of the displaced water. The average density of the person, based on his volume and his dry weight, can be used to determine what percentage of his body is fat.

Triathletes, whose body fat is likely to be very low, demonstrate their appreciation of the principles of buoyancy by preferring to wear wetsuits during swimming events. Lean people tend to sink, and a wetsuit helps an athlete float since it is less dense than water, reducing the energy spent on staying up and allowing more to be spent on moving forward. Because of this effect, triathlons ban wetsuits in warm water events where they are not strictly necessary for survival. You wouldn't want to give those triathletes any breaks before they bike 180 kilometers and then run over 40 kilometers!

· equals weight of the displaced fluid



This chunk of African ironwood weighs 43 N, and it displaces 0.0030 m^3 of water. What is the buoyant force on it?

$m = \rho V$

$m = (1000\ \text{kg/m}^3)(0.0030\ \text{m}^3)$

$m = 3.0\ \text{kg}$

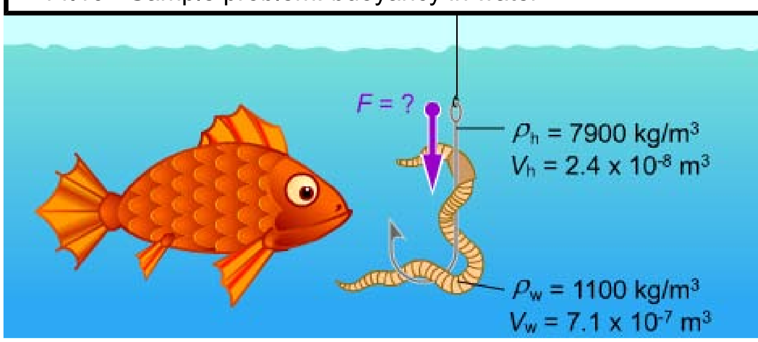
$mg = (3.0\ \text{kg})(9.80\ \text{m/s}^2) = 29\ \text{N}$

$F = 29$ N, directed up

Objects fabricated from materials denser than water can float. A steel boat floats since its hull encloses air, which means the average density of the volume enclosed by the boat is less than that of water. Observe what happens when someone steps into a small boat: It sinks slightly as more water is displaced to balance the person's weight. If the boat is overloaded with cargo, or if water enters the hull, its average density will surpass that of water and the boat will sink (fast-forward to the end of the film *A Perfect Storm* for a graphic example of the latter problem).

Often, the concept of buoyancy is applied to water, but it also applies to other fluids, such as the atmosphere. Blimps and hot air balloons use buoyancy to float in the air. A blimp contains helium, a gas lighter than air. The weight of the air it displaces is greater than the weight of the blimp, so it floats upward until it reaches a region of the atmosphere where the air is less dense and the weight of the blimp equals the weight of the displaced air.

14.10 - Sample problem: buoyancy in water



A stainless steel hook with a worm dangles underwater at the end of a fishing line.

What is the net downward force that the bait combination exerts on the line?

A stainless steel fishing hook with a worm is in the water at the end of a line, dangled with the intent of attracting the wily fish. For the density of the hook we use the density of stainless steel, 7900 kg/m^3.

We will refer to the combination of the hook and worm as "the bait."

Variables

net force of bait on line	F
acceleration of gravity	$g = 9.80\ \text{m/s}^2$
buoyant force on bait combination	F_b

	hook	worm	displaced water
density	$\rho_h = 7900\ \text{kg/m}^3$	$\rho_w = 1100\ \text{kg/m}^3$	$\rho_{H2O} = 1000\ \text{kg/m}^3$
volume	$V_h = 2.4\times10^{-8}\ \text{m}^3$	$V_w = 7.1\times10^{-7}\ \text{m}^3$	V_{H2O}
mass	m_h	m_w	m_{H2O}
weight	$m_h g$	$m_w g$	$m_{H2O} g$

Strategy

1. Calculate the weight of the hook, a downward force.
2. Calculate the weight of the worm, another downward force.
3. From the combined volumes of the hook and worm, compute the volume and the weight of the displaced water. Use this value for the upward buoyant force F_b on the bait combination.
4. With all the contributing downward and upward forces known, calculate the net force exerted by the bait on the line.

Physics principles and equations

Use the definitions of density and weight,

$$\rho = \frac{m}{V}, \qquad \text{weight} = mg$$

Archimedes' principle states that the upward buoyant force on the bait equals the weight of the water it displaces.

Step-by-step solution

Calculate the weight of the hook.

Step	Reason
1. $m_h = \rho_h V_h$	definition of density
2. $m_h g = \rho_h V_h g$	multiply by g
3. $m_h g = (7900\ \frac{\text{kg}}{\text{m}^3})(2.4\times10^{-8}\ \text{m}^3)(9.80\ \frac{\text{m}}{\text{s}^2})$ $m_h g = 1.9\times10^{-3}\ \text{N}$	evaluate

Calculate the weight of the worm.

Step	Reason
4. $m_w g = \rho_w V_w g$	equation 2, for the worm
5. $m_w g = (1100\ \frac{\text{kg}}{\text{m}^3})(7.1\times10^{-7}\ \text{m}^3)(9.80\ \frac{\text{m}}{\text{s}^2})$ $m_w g = 7.7\times10^{-3}\ \text{N}$	evaluate

Calculate the weight of the displaced water, and from that the buoyancy of the bait combination.

Step	Reason
6. $V_{H2O} = V_h + V_w$	add volumes
7. $m_{H2O} g = \rho_{H2O}(V_h + V_w)g$	substitute equation 6 into equation 2
8. $m_{H2O} g = (1000\ \frac{\text{kg}}{\text{m}^3})(7.3\times10^{-7}\ \text{m}^3)(9.80\ \frac{\text{m}}{\text{s}^2})$ $m_{H2O} g = 7.2\times10^{-3}\ \text{N}$	evaluate
9. $F_b = 7.2\times10^{-3}\ \text{N}$	Archimedes' principle

Finally, the net force is the sum of the buoyancy upward and the weight of the bait downward.

Step	Reason
10. $F = F_b + (-m_h g) + (-m_w g)$	net force
11. $F = (7.2\times 10^{-3} - 1.9\times 10^{-3} - 7.7\times 10^{-3})$ N $F = -2.4\times 10^{-3}$ N	evaluate

The negative value of the net force indicates that the bait combination is pulling down on the line.

The quantities involved in this problem are rather small: The weight of the bait, its buoyancy and the net downward force all have magnitudes in the thousandths of newtons. Bait like this will sink, but not very quickly, due to the resistance of water to its motion. For this reason, fishermen often use lead weights called "sinkers" to cause the bait to sink faster to a depth where fish are feeding.

14.11 - Interactive checkpoint: astronaut training

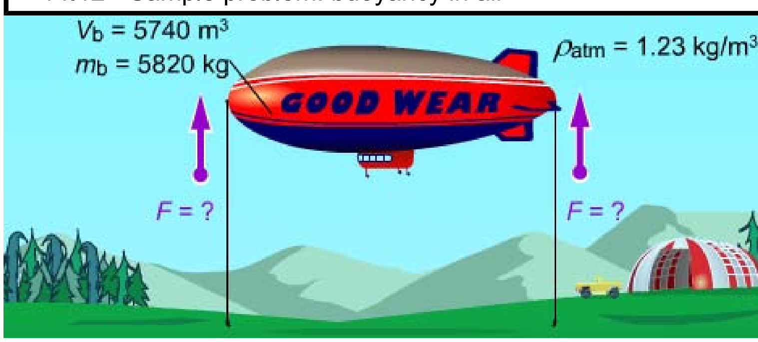
A steel wrench weighs 9.55 N in air. What is the net force on the wrench when it is underwater? The density of steel is 7.90×10^3 kg/m^3 and the density of water is 1.00×10^3 kg/m^3. The downward direction is negative.

Astronauts regularly train for spacewalk maintenance missions by practicing similar tasks underwater in a giant NASA training pool. Floating underwater in this pool partially simulates working in weightless conditions, as you will see by calculating the net force on a steel wrench underwater.

Answer:

$\Sigma F =$ ______ N

14.12 - Sample problem: buoyancy in air



This blimp is tethered to the ground by two vertical mooring lines that are each under the same tension.

What is the net upward force that the blimp exerts on each mooring line?

The blimp you see above is moored at its air station. There are two mooring lines, one at each end of the blimp, to keep it from floating away. Modern airships of this type are filled with helium, whose low density ensures that it has significant buoyancy in the atmosphere. Hydrogen, an even lower density gas, is no longer used because of the dangers associated with its flammability.

Variables

The mass of the blimp is its total inflated mass, including the masses of its envelope, passenger car, engines, and helium gas.

force on one mooring line	F
acceleration of gravity	$g = 9.80\ \text{m/s}^2$
buoyant force on blimp	F_b

	blimp	displaced air
density	ρ_b	$\rho_{atm} = 1.23\ \text{kg/m}^3$
volume	$V_b = 5740\ \text{m}^3$	V_{atm}
mass	$m_b = 5820\ \text{kg}$	m_{atm}
weight	$m_b g$	$m_{atm} g$

Strategy

1. Calculate the downward weight of the blimp.
2. Calculate the volume of the blimp, and then determine the volume and the weight of the displaced air. Use this value for the upward buoyant force F_b on the blimp.
3. With the other forces calculated, compute the net upward force on each mooring line.

Physics principles and equations

Use the definitions of density and weight.

$$\rho = \frac{m}{V}, \quad \text{weight} = mg$$

Archimedes' principle states that the upward buoyant force on the blimp equals the weight of the fluid it displaces.

Step-by-step solution

Calculate the weight of the blimp.

Step	Reason
1. $m_b g = (5820\ \text{kg})(9.80\ \text{m/s}^2)$ $m_b g = 5.70 \times 10^4\ \text{N}$	definition of weight

Calculate the weight of the displaced air, which equals the magnitude of the buoyant force.

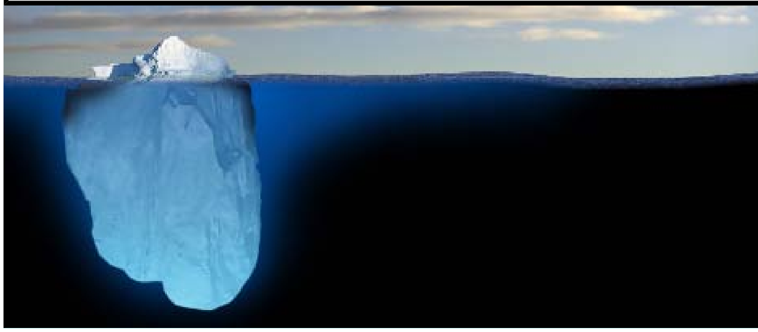
Step	Reason
2. $V_{atm} = V_b$	equal volumes
3. $m_{atm} = \rho_{atm} V_{atm}$	definition of density
4. $m_{atm} = \rho_{atm} V_b$	substitute equation 2 into equation 3
5. $m_{atm} g = \rho_{atm} V_b g$	multiply by g
6. $m_{atm} g = (1.23\ \text{kg/m}^3)(5740\ \text{m}^3)(9.80\ \text{m/s}^2)$ $m_{atm} g = 6.92 \times 10^4\ \text{N}$	evaluate
7. $F_b = 6.92 \times 10^4\ \text{N}$	Archimedes' principle

Now calculate the upward force on each mooring line.

Step	Reason
8. $2F = F_b + (-m_b g)$	net force
9. $F = \dfrac{6.92 \times 10^4 \text{N} - 5.70 \times 10^4\ \text{N}}{2}$ $F = 0.61 \times 10^4\ \text{N} = 6100\ \text{N}$	solve for F and evaluate

The positive value of the force on each mooring line indicates that the blimp is pulling upward on them. The tension on each mooring line translates to about 1400 pounds. This is well within the specifications of a climbing rope that you could buy at any outdoor recreation store.

14.13 - Sample problem: buoyancy of an iceberg



What fraction of this iceberg is submerged below the water?

You may have heard the expression "it's just the tip of the iceberg"; icebergs are infamous for having nine-tenths of their volume submerged below the surface of the sea. The composite photograph above shows just how dangerous icebergs can be for navigation. The submerged portion not only extends downward a great distance, it may also extend sideways to an extent that is not evident from above. A ship might easily strike the submerged portion without passing especially close to the visible ice.

Use the values stated below for the densities of ice and of seawater at $-1.8°\mathrm{C}$, which is the freezing point of seawater in the arctic. The iceberg is in static equilibrium, moving neither up nor down.

Variables

The upward buoyant force on the iceberg is F_b.

	iceberg	displaced water
density	$\rho_{ice} = 917\ \mathrm{kg/m^3}$	$\rho_{H2O} = 1030\ \mathrm{kg/m^3}$
volume	V_{ice}	V_{H2O}
mass	m_{ice}	m_{H2O}

Strategy

1. Use equilibrium to state that the buoyant force acting on the iceberg equals its weight. Archimedes' principle allows you to express the buoyant force in terms of the weight of the displaced water. Replace the equilibrium equation with one stating that the mass of the iceberg equals the mass of the displaced water.
2. Use the definition of density to replace the masses in the previous equation by products of density and volume. Solve for the ratio V_{H2O}/V_{ice}, and evaluate it.

Physics principles and equations

Newton's second law.

$$\Sigma F = ma$$

Archimedes' principle states that the buoyant force on an object in a fluid equals the weight of the fluid it displaces.

The definition of density is

$$\rho = \frac{m}{V}$$

Step-by-step solution

We begin by stating the condition for static equilibrium, and then we apply Archimedes' principle.

Step	Reason
1. $F_b + (-m_{ice}g) = 0$ $F_b = m_{ice}g$	equilibrium
2. $m_{H2O}\,g = m_{ice}\,g$	Archimedes
3. $m_{H2O} = m_{ice}$	simplify

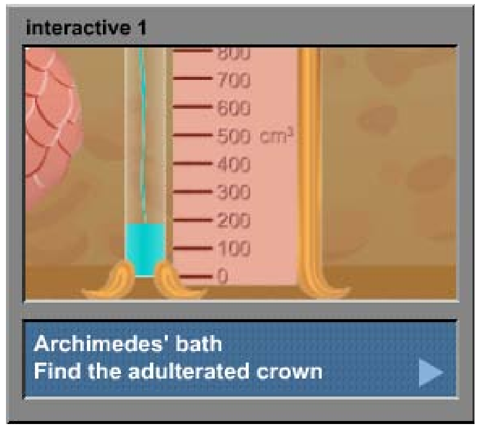
Now we replace the masses in the previous equation by products of densities and volumes, solve for a ratio of volumes, and evaluate.

Step	Reason
4. $\rho_{H2O} V_{H2O} = \rho_{ice} V_{ice}$	definition of density
5. $\frac{V_{H2O}}{V_{ice}} = \frac{\rho_{ice}}{\rho_{H2O}}$	rearrange
6. $\frac{V_{H2O}}{V_{ice}} = \frac{917\ \text{kg/m}^3}{1030\ \text{kg/m}^3} = 89.0\%$	evaluate

Since the volume V_{H2O} of the displaced water equals the volume of the submerged portion of the iceberg, we have shown that 89.0% of the iceberg is submerged.

For most substances, the solid phase is denser than the liquid phase, meaning a solid sinks when immersed in a liquid composed of the same substance. Water is very unusual in this respect: the solid phase (ice) floats in liquid water. This proves crucial to life on Earth for a variety of reasons.

14.14 - Interactive problem: Eureka!



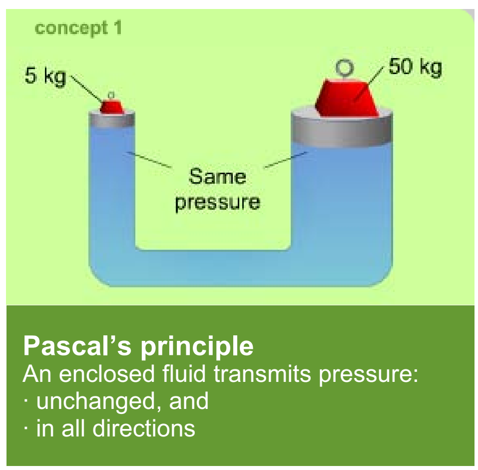
In this simulation you play the role of the ancient Greek mathematician and physicist Archimedes. As the story goes, in olden Syracuse the tyrant Hieron suspected that a wily goldsmith had adulterated one of his kingly crowns during manufacture by adding some copper and zinc to the precious gold. His Royal Highness asked Archimedes to discover whether this was indeed the case.

Archimedes pondered the problem for days. Then, one day, in the public bath, he observed how the water level in the pool rose as he eased himself in for a good soak, and suddenly realized that the answer was right in front of his eyes. Ecstatic, he supposedly leapt from the bath and ran dripping through the streets of the city crying "*Eureka!*" (I have found it!) The adulterated crown was quickly identified, and the goldsmith roundly punished.

Your task is to determine the nature of Archimedes' insight and apply it in the simulation to the right. You have two crowns and a bar of pure gold. You have already observed their masses using a balance scale. You can measure the volume of a crown or bar in the simulation by dragging it to the bath and noting how much water it displaces. You see the tube that measures the displaced liquid in the illustration at the right. Decide which crown has been altered and drag it to the king's palace. The appropriate consequences will be enacted.

14.15 - Pascal's principle

Pascal's principle: **Pressure in a confined fluid is transmitted unchanged to all parts of the fluid and to the containing walls.**



If you jump down on one side of a waterbed, a person sitting on the other side will get a jolt up. This illustrates Pascal's principle: A variation in pressure in the enclosed fluid is being transmitted unchanged throughout the fluid.

In the illustration to the right, you see a five-kilogram mass "balancing" a 50-kilogram mass. How does something that seems so counterintuitive − a small mass balancing a large mass − occur?

First, Pascal's principle asserts that the **pressure** exerted by the weight of the first mass on the fluid is transmitted to the second mass unchanged. The two masses "balance" because the surface area of the fluid under the 50-kilogram mass is 10 times larger than the surface area under the five-kilogram mass. The pressure is the same under both masses, but since the surface area is 10 times larger under the more massive object, the upward force, the pressure times the area, is 10 times greater than it is under the less massive object, so the system is in equilibrium. Although we focus on the pressures supporting the weights, according to Pascal's principle the pressure acts in all directions at every point in the fluid, so it presses on the walls of the hydraulic system as well.

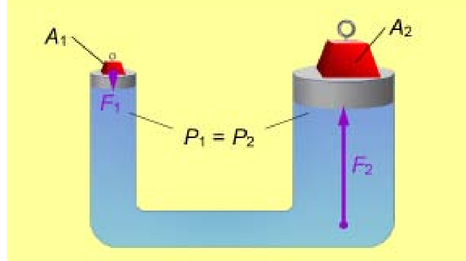
In some ways, the equilibrium in such a system is similar to a small mass located at the end of a long lever arm balancing a large mass located close to the lever's pivot point. In fact, this similarity has caused systems like those shown to the right to be called *hydraulic levers*. You may be familiar with them from the "racks" in car repair shops: they are the gleaming metal pistons that lift cars.

Illustrations of hydraulic levers are shown to the right. In each case the pressure on both pistons must be the same in order to conform to

Pascal's principle, but the **force** will depend on the areas of the pistons. In the example problem you are asked to calculate the downward force needed on the left piston to lift the small automobile on the right. It turns out that this force is about 20 newtons (only 4½ pounds)!

Do you get "something for nothing" when you raise a heavy car with such a small force? No, the amount of work you do equals the work the piston on the right does on the car. You apply less force, but through a greater displacement. If you press your piston down a meter in the scenario shown in Example 1, the car will rise only about 1.6 millimeters, as you can verify by considering the incompressible volume of water displaced on each side.

equation 1

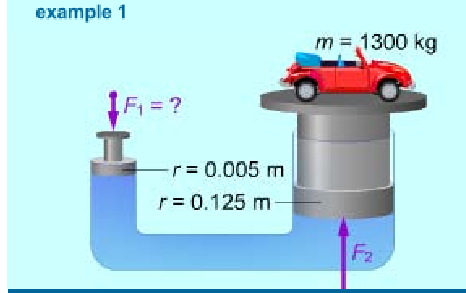


Pascal's principle

$$P_1 = P_2$$
$$F_1/A_1 = F_2/A_2$$

P = pressure
F = force
A = surface area

example 1



How much force must be exerted on the left-hand piston to lift the automobile on the right?

$$P_1 = P_2$$
$$F_1/A_1 = F_2/A_2$$
$$F_1 = mgA_1/A_2$$
$$F_1 = \frac{(1300\ \mathrm{kg})(9.80\ \frac{\mathrm{m}}{\mathrm{s}^2})\pi(0.005\ \mathrm{m})^2}{\pi(0.125\ \mathrm{m})^2}$$
$$F_1 = 20.4\ \mathrm{N}$$

14.16 - Ideal fluid flow

Sir Horace Lamb, a distinguished British mathematician and physicist, once said, "I am an old man now, and when I die and go to Heaven there are two matters on which I hope for enlightenment. One is quantum electrodynamics and the other is the turbulent motion of fluids. And about the former I am really rather optimistic."

The analysis of turbulent fluid flow is quite challenging. We focus instead on *ideal fluid flow* : In order to present the fundamentals of fluid flow, we make a few simplifying assumptions about a fluid:

1. **Steady flow**. The particles making up the fluid move past any fixed point in the flow with a constant speed and direction over time: that is, at a constant velocity. This kind of flow is called *steady flow*, or sometimes *streamline flow*. You might see steady flow in certain parts of a river, for example, even though it is turbulent near its edges or after it passes over obstacles. This is illustrated to the right.
2. **Incompressible**. The fluid is assumed to be *incompressible*. That is, its density is constant. This is a good approximation for liquids, but not a valid assumption for gases.
3. **Nonviscous**. Molasses in January and old motor oil are viscous. These fluids do not flow readily. For them, the process of flowing consumes energy. A *nonviscous* fluid flows easily, with no dissipation of energy. We can treat water as

concept 1

Ideal fluid flow
Simplify fluid flow by assuming:
· constant velocity at any point
· incompressible fluid
· nonviscous fluid
· irrotational flow

a nonviscous fluid; it has little resistance to flow.

4. **No rotation**. In an *irrotational* flow, an object in the fluid will not rotate. The smooth parts of the stream you see to the right could be modeled as regions of irrotational flow. A stick might float down those parts without turning. A stick bobbing below the waterfall is likely to rotate as it encounters eddies in the more complex flow there.

A fluid with the four properties above is called an ideal fluid. The smooth water in the river can be modeled as an ideal fluid.

Other portions of the stream are not ideal fluids. Downstream from points where it encounters obstacles, the water exhibits *unsteady* flow: At fixed points in these regions, the velocities of passing water particles may vary over time. Where you see “white water,” the flow is highly erratic, and the water is displaying *turbulent* flow.

14.17 - Streamline flow

Streamline flow: A fluid flow in which the fluid’s velocity remains constant at any particular point.

concept 1

Streamline flow
At different points, velocities can differ
At any point, velocity constant over time
Image courtesy of Lexus

Steady, or streamline, flow is one of the characteristics of ideal fluid flow. Streamline flow is particularly easy to demonstrate with the flow of a gas, although since gasses are compressible they are not ideal fluids.

At the right, you see one *streamline* traced by smoke in the diagram, flowing around an automobile. As long as the car does not rotate, the streamline stays the same over time. Any particle of the fluid will follow some streamline, visible or not, as it flows past the car. As the car is rotated in the video, you get a chance to see the paths followed by different streamlines of air flowing around various parts of its body.

Engineers use wind tunnels to photograph streamlines. Powerful fans blow air past an object like a car or an airplane, and dyes or smoke are injected into the airflow at several points and carried downstream so that the streamlines are made visible. Engineers analyze the streamlines to investigate the air resistance of a particular car design, or the amount of lift (upward force) generated by an airplane wing.

The velocity of streamline fluid flow can vary from point to point. Air moves past an airplane wing or auto body with different velocities at different points (for example, it moves faster over the tops of these objects than beneath them). The tangent to a streamline at a point coincides with the direction of the velocity vectors of the fluid particles passing by the point.

In streamline flow, the fluid has a constant velocity at all times **at a given point**. The velocity of any given air particle in the visible streamline on the right may change as the flow carries it downstream past the stationary automobile. In particular, a change in the direction of the streamline reflects a change in velocity. However, all the particles in the streamline pass the same point with the same velocity.

How can you conclude that the velocity remains constant at each point? Consider what would happen if the speed of the fluid flow at a point were to vary over time. If the speed increased, the affected particles would collide with particles ahead of them; if it decreased, particles from behind would collide. The resulting collisions would cause an erratic flow, changing the streamline, as would changes in the direction of the particles’ motions. The constancy of the streamlines over time indicates that the velocity at each point does not change.

14.18 - Fluid continuity

Fluid equation of continuity: The volume flow rate of an ideal fluid flowing through a closed system is the same at every point.

concept 1

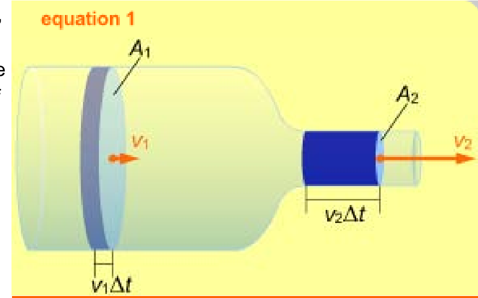
Fluid continuity
Fluid flows past each point at same rate

Turn on a hose and watch the water flow out, and then cover half the hose end with your thumb. The water flows faster through the narrower opening. You have just demonstrated the fluid equation of continuity: How much volume flows per unit time − the volume rate of flow − stays constant in a closed system. The increased speed of the flow at the opening balances the decreased cross sectional area there. Of course after the water leaves the hose with its new speed, it is no longer in the closed system, and you cannot apply the equation of continuity to the resulting spray of droplets. They spread out without slowing down.

The fluid equation of continuity can be observed in rivers, whose courses approximate closed systems. River water flows more quickly through narrow or shallow channels, called rapids, and more slowly where riverbeds are wider and deeper. This relationship is alluded to by the proverb, “Still waters run deep.”

The constancy of fluid flow rate is summarized in the continuity equation that appears as the first line in Equation 1. In this equation, stated for two arbitrary points P_1 and P_2, the rate of flow is measured as the *mass flow rate*. The mass flow rate equals the product of the speed of the fluid, its density, and the cross sectional area it flows through. It is measured in kilograms per second (kg/s). In a closed system (no leaks, no inflows) the mass flow rate is the same past all points.

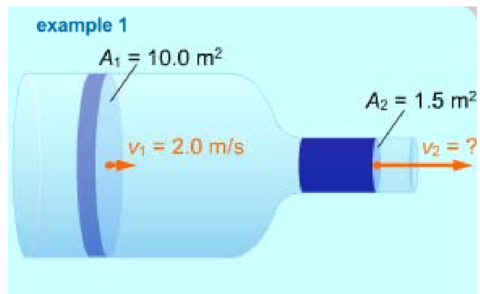
Since we assume that an ideal fluid is incompressible, having a density that is constant, we can cancel the density factor from both sides of the first continuity equation. This enables us to say that the speed of the fluid times its cross sectional area is everywhere the same. This is stated by the second equation, which expresses continuity in terms of the *volume flow rate*, represented by R. The volume flow rate is measured in cubic meters per second (m^3/s). If the cross sectional area decreases (as when the pipe illustrated in Equation 1 narrows), the speed of the fluid flow increases, and R remains the same.



Mass and volume flow rates constant

$$v_1\rho_1A_1 = v_2\rho_2A_2$$
$$v_1A_1 = v_2A_2 = R$$

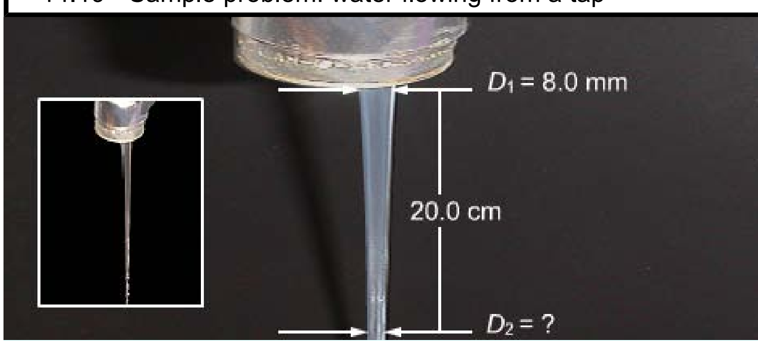
v = speed of fluid
ρ = density of fluid
A = cross-sectional area of flow
Fluid is incompressible
R = volume flow rate



What is the speed of the ideal fluid in the narrow part of the pipe?

$v_1A_1 = v_2A_2$
$v_2 = v_1A_1/A_2$
$v_2 = (2.0\text{ m/s})(10\text{ m}^2) / 1.5\text{ m}^2$
$v_2 = 13\text{ m/s}$

14.19 - Sample problem: water flowing from a tap



Water emerges from the tap at a speed of 0.50 m/s. What is the diameter of the flow 20.0 centimeters below the tap?

The stream of water falling from the tap in the illustration above behaves as an ideal fluid until turbulence sets in, more than 20.0 cm below the tap. We can consider the stream you see as a closed system where the equation of continuity applies. In doing so we will ignore certain other factors, such as surface tension, that become dominant below 20.0 cm and cause the stream to break up into droplets. (Note: We compressed the length of the water column in the close-up photograph above so we could show the change in diameter more clearly.) We assume that the column of water has a circular cross section at all points.

Variables

	at tap	20.0 cm below
diameter of flow	$D_1 = 8.0$ mm	D_2
area of flow	A_1	A_2
speed of flow	$v_1 = 0.50$ m/s	v_2
height	$y_1 = 0$ cm	$y_2 = -20.0$ cm
acceleration of water	$a_y = -9.80$ m/s^2	$a_y = -9.80$ m/s^2

Strategy

1. Find the speed of the water 20.0 cm below the tap.
2. Use the fluid equation of continuity to find the cross sectional area and the diameter of the flow 20.0 cm below the tap.

Physics principles and equations

A motion equation that proves useful,

$$v_2{}^2 = v_1{}^2 + 2a_y\Delta y$$

The fluid equation of continuity,

$$v_1A_1 = v_2A_2$$

Mathematics principle

The area of a circle of diameter D is $A = \pi(D/2)^2$.

Step-by-step solution

First we find the speed of the water after it has fallen 20.0 cm.

Step	Reason
1. $v_2{}^2 = v_1{}^2 + 2a_y\Delta y$	motion equation
2. $v_2^2 = (0.50\ \frac{\text{m}}{\text{s}})^2 + 2(-9.80\ \frac{\text{m}}{\text{s}^2})(-0.200\ \text{m})$	enter values
3. $v_2 = 2.0$ m/s	evaluate

Knowing the flow speed at the second point in the water stream, we can use the continuity equation to find the area and diameter of the flow there.

Step	Reason
4. $v_1A_1 = v_2A_2$	fluid equation of continuity
5. $v_1\pi(D_1/2)^2 = v_2\pi(D_2/2)^2$	area of circle
6. $D_2^2 = \dfrac{v_1 D_1^2}{v_2}$	solve for $D_2{}^2$
7. $D_2^2 = \dfrac{(0.50\ \text{m/s})(0.0080\ \text{m})^2}{2.0\ \text{m/s}}$	enter values
8. $D_2 = 0.0040$ m $= 4.0$ mm	evaluate, take square roots

The diameter of the water column contracts to half its original size after a fall of 20.0 cm. This change is reflected in the photograph above.

14.20 - Bernoulli's equation

Bernoulli's equation applies to ideal fluids. It was developed by the Swiss mathematician and physicist Daniel Bernoulli (1700-1782). The equation is used to analyze fluid flow at different points in a closed system. It states that the sum of the pressure, the KE per unit volume, and the PE per unit volume has a constant value. Concept 1 shows an idealized apparatus for determining these three values at various points in such a system.

A simplified form of Bernoulli's equation is shown in Equation 1. It applies to horizontal flow, in which the PE of the fluid is everywhere the same. In such a system, the sum of just the pressure and the kinetic energy per unit volume is constant. Since the expression for the "KE" uses the density ρ of the fluid in place of mass, it describes energy per unit volume: the *kinetic energy density*.

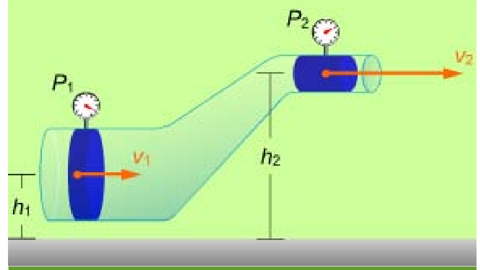
To illustrate the simplified equation, we use the horizontal-flow configuration shown in the diagram of Equation 1. The pressure of the fluid is measured where it passes the gauges. If the speed of the fluid is known at the first gauge, its speed at the second gauge can be calculated using Bernoulli's equation.

When the sum of the pressure and kinetic energy density equals a constant in a system, as in the case of horizontal flow, it is often useful to set the sum of these values at one point equal to the sum of the values at another point. This is stated for points P_1 and P_2 in Equation 2. An implication of the simplified form of Bernoulli's equation is the *Bernoulli effect*: When a fluid flows faster, its pressure decreases.

Airplane wings, like the one shown in Equation 2, take advantage of the Bernoulli effect. Air travels faster over the upper surface of the wing than the lower, because it must traverse a longer path in the same amount of time. A faster fluid is a lower pressure fluid; the result is that there is more pressure below the wing than above. This causes a net force up, which is called *lift*. (The Bernoulli effect is only one way to explain how wings work. Lift can also be explained using Newton's third law in conjunction with a fluid phenomenon called the *Coanda effect*. The topic of wing lift engenders much discussion.)

In Equation 3 you see a general form of Bernoulli's equation, which also accounts for differences in height and the resulting differences in potential energy density. It states that the sum of the pressure and the kinetic energy density, plus the potential energy density, is constant in a closed system. At a higher point in such a system, the potential energy density of the fluid is greater. At that point either the pressure or the kinetic energy density, or both, must be less than they are at lower points.

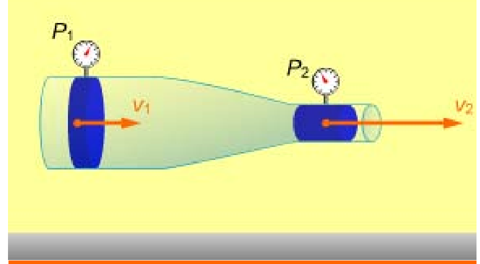
concept 1



Bernoulli's equation

The sum of:
· pressure
· *KE* / unit volume
· *PE* / unit volume
equals a constant in a closed system

equation 1

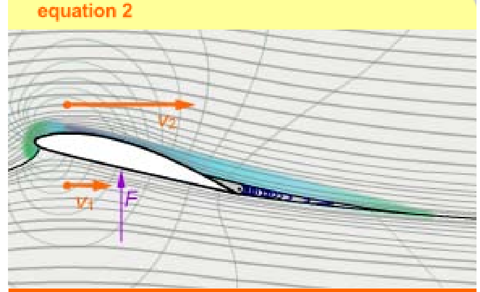


For horizontal flow

$$P + \tfrac{1}{2}\rho v^2 = k$$

P = pressure
ρ = constant density of fluid
v = speed of flow
k = a constant for the system

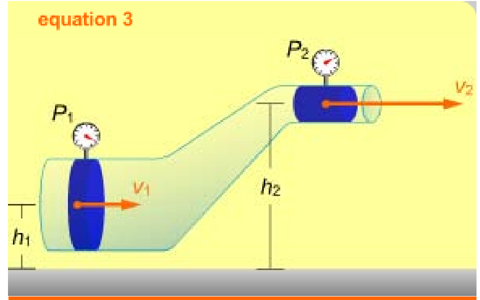
equation 2



The "Bernoulli effect"

$$P_1 + \tfrac{1}{2}\rho v_1^2 = P_2 + \tfrac{1}{2}\rho v_2^2$$

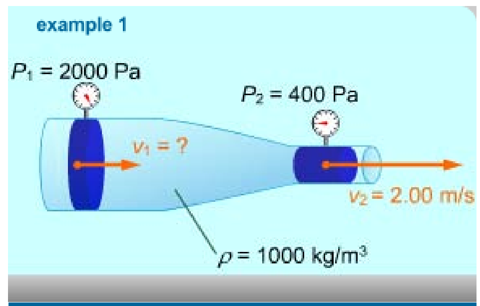
If $v_2 > v_1$, then $P_1 > P_2$
· Net upward force on wing is "lift"



General form of the equation
Includes potential energy density

$$P_1+\frac{1}{2}\rho v_1^2+\rho g h_1 = P_2+\frac{1}{2}\rho v_2^2+\rho g h_2$$

g = acceleration of gravity
h = height



What is the speed of the fluid at point 1?

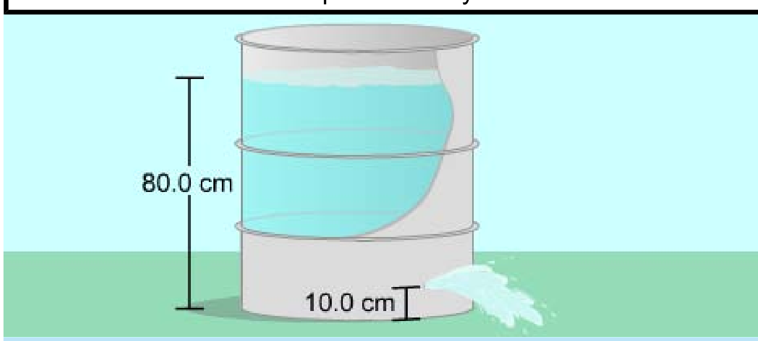
$P_1 + ½\rho v_1^2 = P_2 + ½\rho v_2^2$

$$v_1^2=\frac{2(P_2-P_1)}{\rho}+v_2^2$$

$$v_1^2=\frac{2}{1000}(-1600)+(2.00)^2$$

v_1 = 0.894 m/s

14.21 - Interactive checkpoint: a leaky barrel

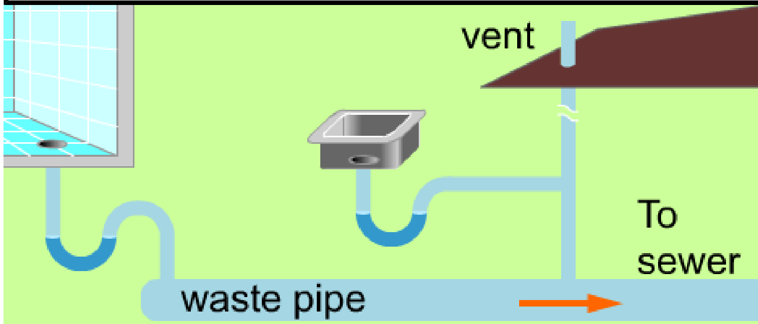


A 55-gallon drum used as a rain barrel is filled with water to a depth of 80.0 cm. A small hole is punctured in the drum 10.0 cm above the bottom. What is the speed of the spouting water? Assume that the drum is large enough that the speed of the sinking surface of the water can be taken to be zero.

Answer:

v_2 = ______ m/s

14.22 - Physics at work: Bernoulli effect and plumbing



Plumbing system with vented sink trap.

concept 1

Drain traps in a plumbing system
Prevent the spread of noxious gases

The Bernoulli effect can help you understand not only the lofty topic of how airplanes fly, but also the more mundane topic of venting in household plumbing systems.

The illustration above shows a shower, a sink, their wastewater pipe, and a “vent” pipe. Water can flow down the shower drain, through a P-shaped “water trap,” and from there to a wastewater pipe that connects to a sewer line. The water trap retains a seal of water that prevents noxious sewer gas from flowing up the waste pipe into the house. The sink in the illustration is likewise connected through a protective P-trap to the waste pipe.

In the illustration, the sink drain is “vented” by an air pipe that extends through the roof of the house. Because it is called a vent pipe, you might think it allows air to leave, but it actually allows air in. Several such vents can be seen on the roofs of most houses. The vents work in conjunction with the water traps that protect the home. What purpose does a vent serve?

concept 2

Traps are vented
To counter Bernoulli effect in wastewater pipes

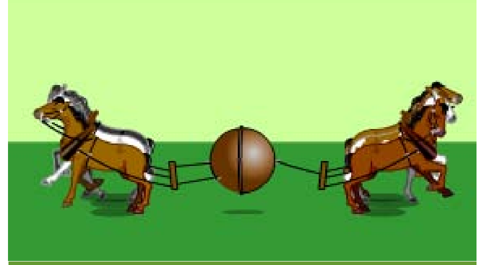
Waste pipes are not pressurized like water supply pipes, and they are always at least slightly inclined so that they ordinarily stand empty. Imagine that someone turns on the shower, causing a stream of water to flow through the waste pipe. As the flow increases, the pressure decreases, in accordance with the Bernoulli effect. If the pressure in the waste pipe were to decrease enough, it could “suck” the water out of the sink’s P-trap, thwarting its protective function.

This is where the vent comes in: When the pressure decreases at the bottom of the sink’s waste connection due to the shower flow, air flows in from the exterior (due to atmospheric pressure), restoring the pressure in the connection. The water remains in place in the trap.

14.23 - The Earth’s atmosphere

The *atmosphere* is the layer of gas, including the oxygen we need in order to survive, surrounding the Earth and bound to it by gravity. The atmosphere receives a great deal of press these days due to environmental concerns such as global warming and ozone holes. Understanding the nature and dynamics of the atmosphere is proving increasingly important to human life.

concept 1



Air pressure demonstration
Near vacuum inside sphere
Horses pull against air pressure
Air pressure can cause great force!

In this section we give a brief overview of three topics relating to the Earth’s atmosphere. The first is the magnitude of atmospheric pressure, and how its existence was first demonstrated. Next, we discuss briefly why there is no hydrogen in the Earth’s atmosphere (unlike the atmosphere of Jupiter). Finally, we provide a general sense of the range of pressures, densities and temperatures of the Earth’s atmosphere at different altitudes.

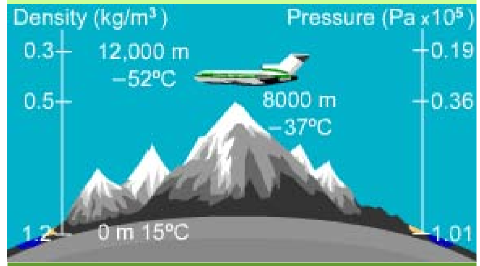
The very existence of atmospheric pressure was once a topic of debate. After all, you do not “feel” air pressure (since the fluid inside your body exerts an equal and opposite pressure) any more than a fish feels water pressure. (At least no fish has ever told us that it feels this pressure.) In a famous demonstration that showed the existence of atmospheric pressure, the German scientist Otto von Guericke created a near vacuum between two copper hemispheres, as shown in Concept 1. Von Guericke, also the mayor of the town of Magdeburg, where he conducted the demonstration in 1654, challenged teams of horses to pull the hemispheres apart. The horses failed: The force of the air pressure was too great for them to overcome.

This clever demonstration of air pressure may seem incredible until you consider that the air pressure on the Earth’s surface is about 101 kilopascals (14.7 pounds per square inch). To get an approximate value for the pressure holding the hemispheres together, we can simplify matters and assume they acted like halves of a cube, one meter on a side. We will further assume that a perfect vacuum was created inside them, so the pressure inside was zero. Each team of horses would then have had to overcome 101,000 N of force, more than 10 tons. Whoa, Nellie!

Von Guericke demonstrated the surprisingly large magnitude of atmospheric pressure. He might have been startled to know how fast the particles that make up the atmosphere move. Sunlight energizes the molecules in the atmosphere, keeping it in a gaseous state, and causes them to fly energetically about, reaching speeds of up to 1600 km/h. The gravitational force of the Earth keeps most of these molecules from flying off into space, but some molecules − especially the lightest ones − do reach escape velocity and leave the planet's atmosphere. This is why there is little or no hydrogen, or any other light gas, in the Earth's atmosphere. On the average, a lighter molecule moves faster than a heavier molecule of the same energy, meaning the lightest molecules most easily reach the 11.2 km/s required to escape the Earth's gravitational pull. The escape velocity is greater on Jupiter, allowing that planet to keep hydrogen in its atmosphere.

The Earth's atmosphere is a gas, which makes it a fluid. However, since it is a gas, its density changes with its height above the planet's surface. The weight of the atmosphere above "compacts" the atoms and molecules below, increasing their concentration (density). We live in an ocean of air, just as fish dwell in an ocean of water. An important difference between the two is that water is relatively incompressible and the ocean has essentially the same density, although different temperatures and pressures, at any depth.

concept 2



At greater altitudes
Pressure, density, temperature decrease

The diagram in Concept 2 shows the Earth's atmosphere schematically. The pressure and density of the atmosphere lessen, as does its temperature, with height. The density of air at sea level and $15°\text{C}$ equals 1.23 kg/m^3; at the higher and chillier altitude of 30,000 meters, where the temperature is $-63°\text{C}$, the density equals 0.092 kg/m^3. At the summit of Mount Everest, a human can breathe barely enough oxygen to stay alive.

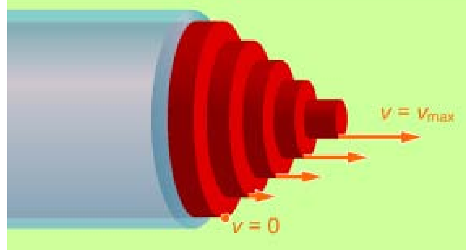
Conversely, in deep diamond mines, air pressure and density both increase, and the temperature rises rapidly with depth. In the deepest mines, it would be impossible for miners to work without the introduction of refrigerated air.

14.24 - Laminar flow

Laminar flow: The flow of layers in a viscous fluid.

Viscous liquids are non-ideal fluids. When a viscous fluid flows, there is internal resistance to the flow. Various "layers" of the fluid flow at different velocities. At the right, you see an illustration of a viscous fluid flowing through a pipe. The outermost layer of the fluid remains stationary, attached to the wall containing it. You can see this occurring in an older bottle of ketchup: A layer of ketchup sticks to the side.

concept 1



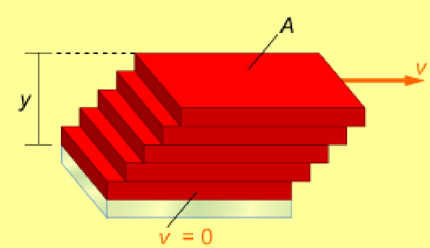
Laminar flow
Layers flow at different speeds
Flow is faster farther from wall

A second layer of fluid flows relatively slowly past this first, fixed layer. Successive layers toward the interior of the pipe flow more quickly, and at the middle of the pipe the innermost "tube" of fluid flows the fastest. This type of flow is called laminar flow. Each thin layer is called a *lamina*.

Returning to ketchup for a moment: Restaurateurs know that a good way to get ketchup to flow is to tap the side of the bottle with a knife, reducing the amount of ketchup bonded along the wall. This often works better than vigorous shaking.

The formula in Equation 1 enables you to calculate the amount of force required to get a layer of ketchup, or other viscous fluid, flowing. The Greek letter η ("eta," rhymes with "wait-uh") represents the *coefficient of viscosity* of the fluid. A is the area of the layer, v the speed at which it is moving, and y the distance of the layer from the container wall. An unusual feature of the equation is that a constant force corresponds not to acceleration, but to a constant velocity. The viscosity of the fluid provides a velocity-dependent force that counters the external force. This may remind you of the analogous concept of terminal velocity in air.

equation 1



Force required to drive layer

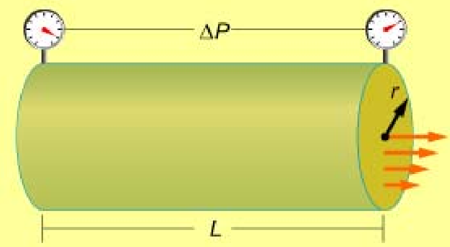
$$F = \frac{\eta A v}{y}$$

F = force
η = coefficient of viscosity
A = area
v = speed of flow
y = distance from container wall
Units of η: poise (P), 0.1 Pa·s

The coefficients of viscosity for liquids tend to be much greater than those for gases. They also vary with temperature. Warming up a fluid causes it to flow more easily, while as the proverb states, molasses flows slowly indeed in the cold month of January. Since an ideal fluid has no viscosity, its viscosity coefficient is zero.

The second formula to the right, called *Poiseuille's law*, is used to calculate the volume flow rate R of a fluid through a cylindrical pipe. This rate increases with the radius r of the pipe and the difference in pressure ΔP between two points separated by a length L of the pipe. R increases proportionally to r^4. In other words, it increases proportionally to the square of the cross sectional **area** of the pipe. It decreases with the coefficient of viscosity and the length of the pipe.

equation 2



Poiseuille's Law

$$R = \frac{\pi r^4 \Delta P}{8\eta L}$$

R = volume flow rate in m^3/s
r = pipe radius
ΔP = pressure difference
η = coefficient of viscosity
L = length

equation 3

	Coefficient of viscosity in mPa·s
Air (0° C)	0.01708
Water (20° C)	1.002
Mercury (25° C)	1.526
Olive oil (20° C)	84.0
Heavy machine oil (15.6° C)	660.6
Corn syrup	2000 - 3000
Chocolate syrup	10,000 - 25,000
Peanut butter	150,000 - 250,000

Viscosity η of various substances

14.25 - Surface tension

Surface tension: A cohesive effect at the surface of a liquid due to the forces between the liquid's atoms or molecules.

Dewdrops on rose petals form tiny spheres.

Above, you see a dewy rose. The drops of water do not spread and flow to cover the petals on which they rest. Instead they contract into tiny spheroids. They do this because the surface tension of water causes each drop to try to minimize its own surface area. The bristles of a wet paintbrush contract into a sleek shape for the same reason. The surface tension of the water on the bristles causes them to pull in. You see this in Concept 1.

Surface tension also enables an insect called a water strider to walk on water. The creature's weight, transmitted to the water's surface through its feet, causes tiny depressions, but the feet do not break through. The surface tension of the deformed liquid surface provides an elastic-like restoring force that balances the insect's weight. A video of this interesting phenomenon is shown in Concept 2.

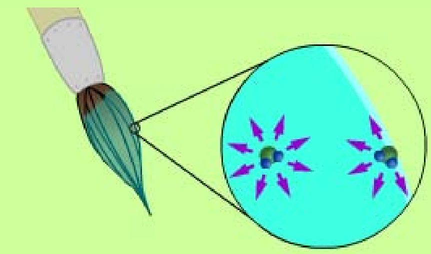
What causes surface tension? In some liquids like water, the molecules that make up the liquid attract each other. These molecules are dipoles; they have regions of positive and negative electrostatic charge. The positive pole of each molecule is attracted to the negative poles of neighboring molecules, and vice versa.

Molecules in the interior of the liquid experience equal attractions in all directions, and so they experience no net intermolecular force. Molecules at the surface of the liquid, however, are pulled on only by the molecules below, and by their neighbors in the surface, so there is a net force pulling them into the interior of the liquid. Because of this, the surface of the liquid tends to contract and consequently to minimize its own area. You see the intermolecular forces illustrated in Concept 1.

In the case of liquid dewdrops, the minimum surface area is roughly spherical. In the case of the water strider, the elastic-like upward force of surface tension on its feet results from the water's tendency to flatten, and so minimize the area, of its surface.

Water has a strong surface tension, but that tension can be reduced. For example, when you heat water, you reduce its surface tension because the faster moving molecules do not attract each other as much as they would at cooler temperatures. Diminished surface tension allows other substances that might be in the water − say, butter − to rise to the surface. Cooks know this (at least implicitly), and they serve soups hot because they will be more flavorful. Adding soap to water also reduces its surface tension − and it can cause a water strider to sink!

concept 1



Surface tension
Contraction of surface of liquid
Due to mutual attraction of molecules

concept 2

Water strider
Surface tension lets insect walk on water
Restoring force balances weight

14.26 - Gotchas

Pressure increases with depth in a fluid. Yes, it does. The farther below the surface of a fluid an object is, the more fluid above it and the greater the pressure on it.

Buoyant force increases with an object's depth in water. Only if more of the object is getting submerged, in which case the buoyant force does increase. But a soda can five meters below the surface, and one 500 meters below the surface, both experience the same buoyant force. This force is equal in magnitude to the weight of the displaced water.

Two surfaces have the same pressure on them, so the pressure must exert the same force on each. No, pressure is force divided by area, so if one surface has a greater area, it experiences a greater force.

Pressure acts in every direction. Yes, this is stated by Pascal's principle. For instance, water pressure pushes both down on the top and up on the bottom of a scuba diver exploring under the sea.

14.27 - Summary

Fluids are substances, liquids and gases, that can flow and conform to the shape of the container that holds them.

A material's density is the amount of mass it contains, per unit volume. Density is represented by the Greek letter ρ and has units of kg/m^3. Unless otherwise stated, substances are assumed to be of uniform density, which means they have the same density at all points.

Pressure is the amount of force on a surface per unit area. The unit of pressure is the pascal (Pa), which is equal to 1 N/m^2.

Fluids can exert forces, and therefore pressure, just as solid objects can. For an object immersed in a liquid, the pressure is the product of the liquid's density, the acceleration of gravity g, and the object's depth. Gauge pressure is the pressure due solely to the liquid, while absolute pressure is the pressure of the liquid plus atmospheric pressure.

Buoyancy is the upward force that results when an object is placed in a fluid, the force that causes a ship to float. Archimedes' principle states that the magnitude of the buoyant force is equal to the weight of the fluid displaced by the object.

Pascal's principle applies to confined fluids. It states that an enclosed fluid will transmit pressure unchanged in all directions.

To simplify the study of fluid flow, we often assume an ideal fluid flow. This means that the flow is streamline flow − it has a constant velocity at every fixed point − and that it is irrotational. The fluid is also assumed to be incompressible and nonviscous.

One property of ideal fluid flow is stated by the fluid equation of continuity. The amount of fluid flowing past every point in a closed system is the same. In other words, the volume flow rate is constant regardless of the size of the area through which the fluid flows.

Another property of ideal fluid flow is described by Bernoulli's equation. The sum of the pressure, kinetic energy density, and potential energy density is constant in a closed system. For horizontal flow, the faster a fluid flows, the lower its pressure. This is called the Bernoulli effect.

One type of non-ideal fluid flow is laminar flow. It occurs when layers of viscous fluid flow at different speeds. In a container or pipe, the layers farther from the wall flow faster. Poiseuille's law gives the volume flow rate of laminar flow through a pipe.

The atmosphere exerts pressure because it is a fluid. But since it is a gas, air density and pressure decrease noticeably with altitude. Despite often being ignored in day-to-day life, air pressure is actually (and demonstrably) quite large.

Surface tension is an effect seen with certain liquids such as water. Polar molecules on the surface of the liquid are attracted toward the interior by unbalanced intermolecular forces, causing the surface to contract, and consequently to minimize its own area and exhibit some elastic-like properties.

Equations

Definition of density

$\rho = m/V$

Definition of pressure

$P = F/A$

Pressure of a liquid

$P = \rho g h$

Pascal's principle

$P_1 = P_2$

$F_1/A_1 = F_2/A_2$

Fluid equation of continuity

$v_1 \rho A_1 = v_2 \rho A_2$

$v_1 A_1 = v_2 A_2 = R$

Bernoulli's equation

For horizontal flow: $P + \frac{1}{2}\rho v^2 = k$

$P_1 + \frac{1}{2}\rho v_1^2 = P_2 + \frac{1}{2}\rho v_2^2$

General form:

$$P_1 + \tfrac{1}{2}\rho v_1^2 + \rho g h_1 = P_2 + \tfrac{1}{2}\rho v_2^2 + \rho g h_2$$

Laminar flow

$$F = \frac{\eta A v}{y}$$

$$R = \frac{\pi r^4 \Delta P}{8 \eta L}$$

Chapter 14 Problems

Chapter Assumptions

Unless stated otherwise, use the following values:

Atmospheric pressure at the Earth's surface: $P_{\text{atm}} = 1.013\times10^5$ Pa

Density of pure water = 1000 kg/m^3

Density of seawater = 1030 kg/m^3

“Standard temperature and pressure” means 0°C and the atmospheric pressure stated above.

Conceptual Problems

C.1 You are standing in your driveway. You measure the pressure inside a bicycle tire with your tire gauge and get a reading of 60 psi. You then don your space suit, take the tire into outer space and repeat the measurement, this time getting a reading of 75 psi. Air was neither added to nor removed from the tire, and its temperature did not change. What explains this discrepancy?

C.2 Scuba divers are instructed to exhale slowly but continuously as they rise to the surface in an emergency situation (such as losing a tank). How is it possible for them to do this?

C.3 If an astronaut took a full bottle of water and ejected it from the pressurized interior of the International Space Station out into space, what would happen? Your friend claims that there would be a sudden overpressure of almost 15 psi from inside the bottle, and that it would expand and explode violently. Do you agree? Explain your answer.

- ◯ Yes. The bottle will explode.
- ◯ No. The bottle will not explode.

C.4 Lurid science fiction stories sometimes dramatize a deep-space event known as "explosive decompression": The villain ejects an innocent victim, without a spacesuit, from a spaceship's airlock, and the victim's eyes bug out and then he or she explodes. Such an event is enacted on the nearly airless surface of Mars in the Schwarzenegger film *Total Recall*. These scenes are inaccurate, but if you are suddenly ejected into space, you should be concerned about the danger of a sudden expansion of a substance in your body. What is this substance?

C.5 A barometer is constructed by inverting a closed-end tube full of mercury and submerging its open end in a mercury reservoir that is open to the air. The mercury in the tube will sink, leaving a vacuum at the closed end, until the system reaches equilibrium with the atmospheric pressure outside the reservoir. For a given atmospheric pressure, does the height of the column depend on the diameter of the tube? Explain.

◯ Yes ◯ No

C.6 A barometer is constructed using a closed-end tube containing a vacuum above a column of mercury, as described in the textbook. On a certain day, the pressure exerted solely by this column of mercury at the bottom of the tube is 1.024×10^5 Pa. Check each of the following quantities that are equal to this measurement.

- ☐ Absolute pressure at bottom
- ☐ Atmospheric pressure

C.7 In a legendary and probably apocryphal story, a physics professor poses a question on a test, "How would you use a barometer to determine the height of a tall building?" In the story, a brilliant but rebellious physics student artfully avoids giving the "correct" answer but gives instead a long list of plausible alternative answers, including the following...

The kinematic answer: I would drop the barometer from the top of the building and time its fall. The equation $\Delta y = v_i t + (1/2)at^2$ would then tell me the building's height.
The pendulum answer: I would tie the barometer to a long string, lift it slightly above the ground, and swing it from the top of the building. The equation $T = 2\pi\sqrt{L/g}$ would then tell me the building's height.
The geometric answer: On a sunny day, I would measure the height of the barometer, the length of its shadow, and the length of the building's shadow. I would then use similar triangles to compute the building's height.
The human-engineering answer: I would go to the building manager and say, "I have here a fine scientific instrument that I will give to you if you tell me the building's height!"

What answer was the professor really looking for?

C.8 An empty boat is placed in a freshwater lake and a mark is painted on the hull at the waterline, a line corresponding to the surface of the water when the vessel is floating upright. The same boat is then transported to Jupiter, and placed into a pool of fresh water that has been prepared just for this comparison experiment. The acceleration due to gravity on Jupiter is 2.6 times what it is on Earth. The new waterline is noted on Jupiter. Compared to the waterline mark on Earth, where is the new waterline mark located on the hull of the boat? Ignore atmospheric effects.

The new waterline mark is
i. higher
ii. lower
iii. at the same place
on the hull.

C.9 An empty boat is placed in a freshwater lake and a mark is painted on the hull at the waterline, a line corresponding to the surface of the water when the vessel is floating upright. The boat is then transported to the Dead Sea, where the liquid density is about 1.2 times that of fresh water due to the high concentration of salts. A waterline mark is noted in the Dead Sea. Compared to the first waterline mark, where is the new waterline mark located on the hull of the boat?

The new waterline mark is
i. higher
ii. lower
iii. at the same place
on the hull.

C.10 You hold a ping-pong ball and a steel ball bearing of the same diameter so that they are submerged underwater. Which one experiences the greater buoyant force? Explain your answer.

i. The ping-pong ball
ii. The ball bearing
iii. The buoyant forces are the same
.

C.11 A ping-pong ball and a steel ball bearing of the same diameter are thrown into a swimming pool. The ball floats, while the bearing sinks. Which one experiences the greater buoyant force? Explain your answer.

i. The ping-pong ball
ii. The ball bearing
iii. The buoyant forces are the same

C.12 A boat carrying a load of bricks is floating in a canal lock. One of the crew members throws a brick overboard, and it sinks. Does the level of the water in the lock rise or fall? Explain your answer.

○ It rises ○ It falls

C.13 Explain how you would float a battleship in a thimbleful of water. This is a real question, not a trick question, except that you have to make one idealizing assumption: Water does not consist of discrete molecules, but is a "fluid" at every scale of magnitude.

C.14 The string on a helium balloon is tied to the floor of a car. The car then makes a right turn. While the car is turning, how does the balloon sway according to observers inside the car? The windows are closed so there is no connection to the air outside.

The balloon
i. sways to the left
ii. stays upright throughout the turn
iii. sways to the right

C.15 Susie is out fishing on a calm, sunny day. She hooks an old tire that is resting on the lake bottom, and hauls it into her boat. In principle, how does the water level of the lake change?

The lake's water level
i. drops very slightly
ii. does not change
iii. rises very slightly

C.16 Susie is out fishing again on another calm, sunny day. She sees a chunk of firewood floating in the water, and hauls it into her boat. In principle, how does the water level of the lake change?

The lake's water level
i. drops
ii. does not change
iii. rises

C.17 Here is a trick that is often performed as a physics demonstration. A vacuum cleaner hose and wand are attached to the "back end" of a canister-type vacuum cleaner so that the wand directs a fountain of air straight upward. Then a light ball is placed on top of the fountain where it bobs around and even tumbles to the side, but always returns to the top of the fountain without actually falling off. Explain how this trick works.

C.18 People are warned not to stand near open doors on airplanes in flight, because they can get "sucked" out of the door. Explain how this might happen.

Section Problems

Section 2 - Density

2.1 Calculate the average population density in people/km^2 for each of the following geopolitical entities. (a) The United States, whose population is 293×10^6 people, and land area is 9.37×10^6 km^2. (b) The world, whose population is 6356×10^6 people, and land area is 149×10^6 km^2. (c) Siberia, whose population is 25.1×10^6 people, and land area is 13.5×10^6 km^2. (d) Hong Kong, whose population is 6.70×10^6 people, and land area is 1098 km^2.

(a) __________ people/km^2
(b) __________ people/km^2
(c) __________ people/km^2
(d) __________ people/km^2

2.2 The density of air at standard atmospheric pressure and 15°C is 1.23 kg/m^3. (a) What is the total mass of the air in a rectangular room that measures 5.25 m × 4.20 m × 2.15 m? (b) What is the weight of the air in the room? (c) What is the weight of the air in the Pentagon office building in Washington, DC, which has a volume of 2.2×10^6 m^3?

(a) __________ kg
(b) __________ N
(c) __________ N

2.3 If you have ever toured a facility (such as a cyclotron laboratory) where people have to be protected against radiation, there may have been lead bricks lying around, and you may have been given the opportunity to heft one. It is a surprising experience. The dimensions of a standard brick are 9.2 cm × 5.7 cm × 20 cm, and the density of lead is 11,300 kg/m^3. (a) What is the mass of a (standard) lead brick? (b) What is the weight of the brick, in pounds?

(a) __________ kg
(b) __________ lb

2.4 In the opening sequence of the movie, *Raiders of the Lost Ark*, the intrepid explorer Indiana Jones deftly swipes a golden idol from a Mayan temple, instantly replacing it with a bag of sand of the same size to neutralize the ancient weight-based booby trap protecting it. (a) The density of gold is 19,300 kg/m$^{3.}$ If the idol's mass is 11.3 kg and it is solid gold, calculate its volume. (b) The density of silica sand is 1220 kg/m^3. What is the mass of a sandbag of equal volume? (c) Is the scene realistic?

(a) __________ m^3
(b) __________ kg
(c) ◯ Yes ◯ No

Section 3 - Pressure

3.1 (a) A fashionable spike heel has an area of 0.878 cm^2. When a 61.4 kg woman walking in this shoe sets her full weight down on the heel, what is the pressure it exerts on the floor? (b) The heel of a "sensible" shoe has an area of 38.3 cm^2. When the same woman sets her weight on this heel, what is the pressure?

(a) __________ Pa
(b) __________ Pa

3.2 During the first half of the twentieth century, research into the behavior of materials at ultrahigh pressures was conducted using huge hydraulic presses. In 1958 scientists had a sudden insight: instead of building ever larger presses to exert ever larger forces on material samples of a given area, why not use relatively modest forces and an extremely small area? They invented the *diamond anvil pressure cell*. This device incorporates two opposed diamonds (the "anvils") whose sharp points have been slightly filed off to produce roughly circular flat surfaces with a diameter of 250 μm. A minute sample of material is placed between these surfaces and the diamonds are pressed together using leverage. (a) If the diamonds are 2.15 cm from the pivot point of the lever and a scientist bears down with a force of 255 N at a distance of 14.2 cm from the pivot point, what force presses the diamonds together? (b) What is the pressure exerted on the sample?

(a) __________ N
(b) __________ Pa

3.3 A 106 kg man is traveling through the snow. He finds that when he wears only his boots, the sole of each of which has an area of 136 cm^2, he sinks into the snow up to his calves. Figuring that he will never get to Grandma's house this way, he dons his snowshoes, made of a lightweight flat material with no holes. Each snowshoe can be modeled as a rectangle, 37.0 cm wide and 115 cm long, with a half-circle added to the front and the back as a "toe" and a "heel." (a) What pressure does the man exert on the frozen crust of the snow when he puts his entire weight on one booted foot? (b) What pressure does the man exert when he puts his weight on one snowshoe?

(a) __________ Pa
(b) __________ Pa

3.4 An automobile has four tires, each one inflated to a gauge pressure of 24.0 psi, or 1.66×10^5 Pa. Each tire is slightly flattened by its contact with the ground, so that the area of contact is 17.5 cm by 12.0 cm. What is the weight of the automobile?

__________ N

3.5 An advertisement for a certain portable vacuum cleaner shows off its power with a photograph of the vacuum-cleaner wand suspending a bowling ball by the strength of its suction. The vacuum cleaner can maintain a moderate vacuum inside the apparatus at an absolute pressure of 3.53×10^4 Pa (against an outside atmospheric pressure of 1.01×10^5 Pa) when the intake wand is closed. The wand is a hollow metal cylinder with an inside diameter of 3.19 cm. What is the weight of the heaviest ball the vacuum cleaner can lift?

__________ N

3.6 Water is generally said to be nearly incompressible. The deepest part of the ocean abyss lies at the bottom of the Marianas trench off the Philippines, at a depth of nearly eleven kilometers. At a depth of 10.0 km, the measured water pressure is an incredible 103 MPa (that's **mega**pascals). (a) If the density of seawater is 1030 kg/m^3 at the surface of the ocean, and its bulk modulus is $B = 2.34 \times 10^9$ N/m^2, what is its density at a depth of 10.0 km? (Hint: Use the *volume stress* equation from the study of elasticity, $\Delta P = -B\,(\Delta V/V_i)$, where the object undergoing the stress has an initial volume V_i, and experiences a change in volume, ΔV, when the pressure changes by ΔP.) (b) By what factor does the density increase?

(a) __________ kg/m^3
(b) __________

Section 4 - Pressure and fluids

4.1 Seawater has a density of 1030 kg/m^3. The Marianas Trench is a deep undersea canyon in the Pacific Ocean off the Philippines. Assuming the seawater is incompressible, and ignoring the contribution of atmospheric pressure, what is the pressure in this trench (a) at a depth of 1.00 km? (b) at a depth of 5.00 km? (c) at a depth of 10.0 km? (Empirically measured pressures are a little larger than those given by these calculations because seawater compresses slightly at great depths.)

(a) __________ Pa
(b) __________ Pa
(c) __________ Pa

4.2 A photograph in the text shows how a Styrofoam® cup gets crushed by great pressure deep under the surface of the sea. Before the cup was crushed, experimenters used colored pens to write data on it, including the absolute pressure (3288 psi) at the depth to which they planned to submerge the cup. (You can inspect this data by right-clicking at a spot on the photograph and selecting Zoom In from the popup menu that appears.) At what depth below the ocean's surface is the pressure equal to 3288 psi? Use the value 1030 kg/m^3 for the density of seawater.

__________ m

4.3 (a) The gauge pressure at the bottom of a particular column of water, open at the top, is equal to 1.013×10^5 Pa, which is atmospheric pressure. How tall is this column? (b) The gauge pressure at the bottom of a particular column of mercury, open at the top, is also equal to P_{atm}. How tall is the mercury column? (c) Convert your answer to part b from meters into millimeters. (d) Convert your answer to part b from meters into inches. (Meteorologists often express variations in atmospheric pressure in terms of "inches" or "millimeters" of mercury.)

(a) __________ m
(b) __________ m
(c) __________ mm
(d) __________ in

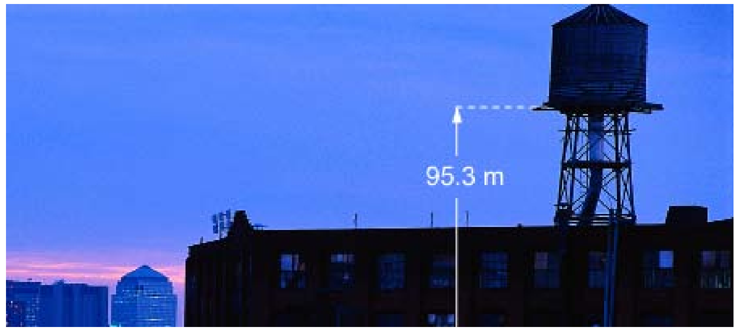
4.4 A creamy salad dressing is made up of heavy cream, corn oil, and vinegar (as well as a pinch of dry mustard, salt and pepper). You put the cream, oil, and vinegar in a glass jar. The density of the cream is 994 kg/m^3, the oil 880 kg/m^3, and the vinegar 1000 kg/m^3. The salad dressing liquids separate into three layers. (a) What is the order of the layers from top to bottom? (b) If the cream layer is 2.20×10^{-2} m tall, the oil layer 2.80×10^{-2} m tall, and the vinegar layer 1.60e−2 m tall, what is the gauge pressure at the bottom of the jar?

(a)
i. Cream/oil/vinegar
ii. Oil/cream/vinegar
iii. Oil/vinegar/cream
iv. Vinegar/cream/oil
v. Vinegar/oil/cream

(b) __________ Pa

4.5 To supply the plumbing system of a New York office building, water needs to be pumped to a tank on the roof, where its height will provide a "head" of pressure for all the floors. The vertical height between the basement pump and the level of the water in the tank is 95.3 m. What gauge pressure does the pump have to apply to the water to get it up to the tank?

__________ Pa



Section 6 - Sample problem: pressure at the bottom of a lake

6.1 In the movie *Creature from the Black Lagoon*, the depth of the freshwater lagoon at its muddy and inscrutable bottom where the Creature lurks is 15.2 m. (a) What is the gauge pressure at the bottom of the lagoon? (b) What is the absolute pressure at the bottom of the lagoon? (c) Who played the Creature in this classic 1954 horror film?

(a) __________ Pa
(b) __________ Pa
(c) i. Ben Chapman
ii. Johnny Depp
iii. Groucho Marx
iv. Fay Wray

6.2 Saturn's moon Titan is the largest moon in the solar system. Its mostly-nitrogen atmosphere exerts a pressure of 1.60×10^5 Pa at the moon's surface (a pressure about 60% greater than P_{atm} at the surface of the Earth). Scientists speculate that it may have lakes and seas of "gasoline," a mixture of liquid methane and ethane. Suppose that there is a gasoline lagoon on Titan, 43.7 m deep, and consisting of 63.7% methane and 36.3% ethane, which have densities of 465 kg/m^3 and 570 kg/m^3 respectively. (a) Assuming that the acceleration due to gravity on Titan is 1.35 m/s^2, what is the gauge pressure at the bottom of the lagoon? (b) What is the absolute pressure at the bottom of the lagoon?

(a) __________ Pa
(b) __________ Pa

Section 7 - Physics at work: measuring pressure

7.1 The gauge pressure at the bottom of a pint of beer, at a depth of 14.6 cm, is 1450 Pa. (a) What is the density of the beer? (b) What is the absolute pressure at the bottom of the pint?

(a) __________ kg/m^3
(b) __________ Pa

7.2 (a) Convert 1.05 bar to torr, using the relationships 1 bar = 10^5 Pa and 1 atm = 1.013×10^5 Pa = 760 torr. Express your answer to the nearest torr. (b) There is a photograph of a dial-type barometer in the textbook, calibrated in both millibars and millimeters of mercury. On this barometer's scales, 1000 millibars (1 bar) appears to be approximately equal to 750 mm (750 torr) of mercury. Is this approximation good to the nearest torr?

(a) __________ torr
(b) ○ Yes ○ No

7.3 The mercury column of an open tube manometer has a height of 0.0690 m. What is the absolute pressure inside the manometer vessel?

__________ Pa

7.4 The pressure inside a manometer vessel is 1.42e+5 Pa. How high is the column of mercury in the open tube of the manometer?

__________ m

Section 9 - Archimedes' principle

9.1 The sizes of ships are commonly expressed in "tons displaced". If a naval vessel displaces 7.67e+3 tons, this means it displaces that weight of water. What is the volume of the water displaced by this ship? Use the fact that 1 ton is equal to 8.90×10^3 N, and take the density of seawater to be 1030 kg/m^3.

__________ m^3

9.2 A cylindrical metal can with a radius of 4.25 cm is floating upright in water. A rock is placed in the can, causing it to sink 0.0232 m deeper into the water. What is the weight of the rock?

__________ N

9.3 The dry weight of a man is 845 N. When he is submerged underwater, the reading on the scale is only 43.2 N. (a) What is the volume of the man? (b) What is the density of the man? (c) The density of fat is 209 kg/m^3, and the density of the "lean" parts of the human body (muscle, bones, and so on) is 2670 kg/m^3. What percentage of the man's body **mass** is "lean body mass"?

(a) __________ m^3
(b) __________ kg/m^3
(c) __________ %

Section 10 - Sample problem: buoyancy in water

10.1 You are fishing off a bridge. Suddenly you feel a tug on the vertical fishing line. Elated, you begin hauling in your catch at constant speed. The creature rears its head above the water and it is...a rubber tire! (a) If the tire is made entirely of hard rubber, with volume 6800 cm^3, and density 1190 kg/m^3, then what is the tension on your fishing line **after** you pull the tire out of the water? Assume that the tire is made entirely of rubber, it is a tire (not an inner tube), and it is punctured so you are not pulling up any water. (b) What is the tension on your fishing line **before** you pull the tire out of the water? Ignore any drag forces from the water.

(a) __________ N
(b) __________ N

10.2 You are fishing off a bridge and feel a tug on the vertical line. This time, your lucky catch is an old boot. (a) Assume that the boot is not punctured, so that as you lift it out of the water at constant speed, you haul up one bootful, or 7500 cm^3, of water along with the boot. If the neoprene rubber making up the boot has volume 435 cm^3 and density 1240 kg/m^3, then what is the tension on your fishing line **after** you pull the boot out of the water? (b) What is the tension in your fishing line **before** you pull the boot out of the water? Ignore any drag forces.

(a) __________ N
(b) __________ N

10.3 According to official US Mint specifications, a quarter-dollar coin has a diameter of 24.26 mm and a thickness of 1.750 mm. It is composed of a copper core, clad with nickel on both faces (as you can see for yourself by examining the edge). 8.330% of the coin, by mass, is nickel. What is the weight of a quarter underwater, to four significant figures? The density of copper is 8960 kg/m^3, and the density of nickel is 8902 kg/m^3.

__________ N

Section 12 - Sample problem: buoyancy in air

12.1 In the motion picture *Danny Deckchair*, based on an actual event, a man attaches 42 helium-filled weather balloons to an aluminum deck chair, steps in, and takes off to experience a series of adventures. (Don't try this at home - or anywhere else!) The weight of the man plus the chair plus the balloons is 927 N. Each balloon is a sphere 1.60 meters in diameter. The density of air at sea level and 15°C is 1.23 kg/m^3. What is the net upward force on Danny and his vehicle right after he leaves the ground? Ignore the volume of the man and of the deckchair.

__________ N

12.2 The Von Hindenburg airship was a famous lighter-than-air vehicle that met a fiery end at its moorage in Lakehurst, New Jersey, in 1937. Called a "dirigible" because, unlike traditional hot-air balloons, it was "directable," the craft used hydrogen, the lightest gas, for buoyancy. The hydrogen capacity of the craft was 200,000 m^3, and together with its cabin, access spaces, and navigation surfaces it displaced 205,000 m^3 of air. Its fully-loaded weight, including hydrogen gas, engines, diesel fuel, crew, 72 passengers, and 76 kg of luxury items such as caviar, was 1.95×10^5 kg. When the ship was safely docked, it was held down by two angled lines, each 20.6 meters long, at a height of 14.2 m above their attachment points on two mooring towers. What was the tension on each mooring line? Assume that the density of air is 1.23 kg/m^3. (Of course, hydrogen is not only the lightest gas, it is also highly flammable, which makes it risky to use in airships.)

__________ N

12.3 In this problem, your task is to analyze a lead balloon. The balloon in question is a hollow sphere of diameter of 20.0 m, constructed from a lead foil 0.25 mm thick and filled with helium gas. Below the balloon, a passenger basket is suspended. The mass of the passenger basket along with the balloon's riggings is 295 kg. For various parts of this problem, you will need to know the density of lead (11,300 kg/m^3), and the densities of air and helium at standard atmospheric temperature and pressure (1.28 kg/m^3 and 0.179 kg/m^3). (a) What is the weight of the lead balloon, including helium, lead, basket and rigging? (b) What is the volume of the air displaced by the balloon? (Ignore any air displaced by the basket and rigging.) (c) What is the buoyant force on the balloon? (d) The maximum "payload" of the balloon, the weight of the passengers and equipment it can lift, equals the net upward force on an unloaded balloon. Can this balloon lift a payload of 7.5×10^3 N?

(a) __________ N
(b) __________ m^3
(c) __________ N
(d) ○ Yes ○ No

Section 13 - Sample problem: buoyancy of an iceberg

13.1 A shipwrecked mariner is stranded on a desert island. He seals a plea for rescue in a 1.00 liter bottle, corks it up, and throws it into the sea. If the mass of the bottle, plus the message and the air inside, is 0.451 kg, what percentage of the volume of the bottle is submerged as it bobs away? Take the density of seawater to be 1030 kg/m^3. For simplicity, assume the bottle and its contents have a uniform density.

__________ %

13.2 (a) A block of balsa wood is placed in water. The density of the wood is 125 kg/m^3. What percentage of the block is submerged? (b) A block of maple wood is placed in water. The density of the wood is 683 kg/m^3. What percentage of the block is submerged? (c) A block of ebony wood is placed in water. The density of the wood is 1200 kg/m^3. What percentage of the block is submerged?

(a) __________ %
(b) __________ %
(c) __________ %

Section 14 - Interactive problem: Eureka!

14.1 Use the information given in the interactive problem in this section to determine which crown is not made of pure gold. Confirm your answer by using the simulation.

○ The 3.20 kg crown
○ The 2.70 kg crown

Section 15 - Pascal’s principle

15.1 An automobile having a mass of 1750 kg is placed on a hydraulic lift in a garage. The piston lifting the car is 0.246 m in diameter. A mechanic attaches a pumping mechanism to a much smaller piston, 1.50 cm in diameter, which is connected by hydraulic lines to the lift. She pumps the handle up and down, slowly lifting the car. What is the force exerted on the small piston during each downward stroke?

__________ N

Section 18 - Fluid continuity

18.1 An incompressible fluid flows through a circular pipe at a speed of 15.0 m/s. The radius of the pipe is 5.00 cm. There is a constriction of the pipe where the radius is only 3.20 cm. How fast must the fluid flow through the constricted region?

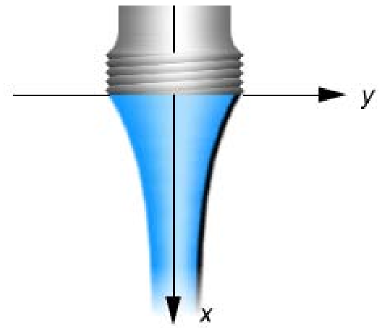
__________ m/s

18.2 The open end of a garden hose is directed horizontally, at a height of 1.25 m above the ground. Water issues from the hose and follows a falling parabolic trajectory to strike the ground 2.41 m away. A gardener holding the hose wishes to water some plants that are 5.12 m distant. What fraction of the hose end should she cover with her thumb? Assume that she continues to hold the hose end horizontally at the same height, and be careful to tell the fraction **covered**, not the fraction left open.

____

Section 19 - Sample problem: water flowing from a tap

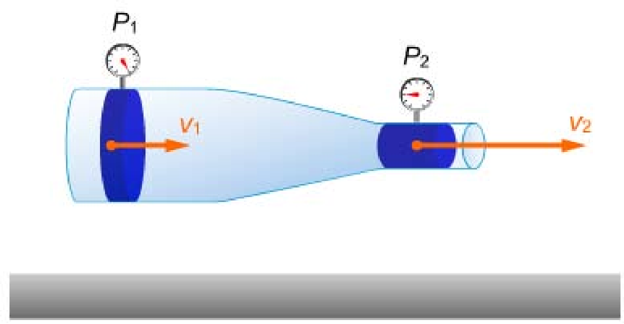
19.1 The illustration shows water flowing from a tap. The diameter of the end of the tap is $2R$, and the water emerges from the tap at speed v_i, subsequently falling under the acceleration due to gravity. The right-hand side of the water column is emphasized with a dark curve. Using an analysis similar to the one in the textbook, write an equation for this curve in the form $y = f(x)$. Note the orientations of the x and y axes.



Section 20 - Bernoulli's equation

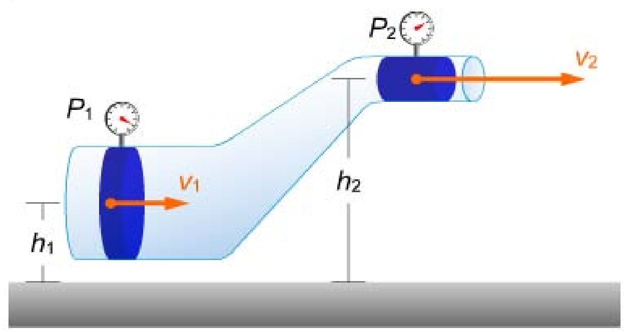
20.1 A stream of water is flowing through the horizontal configuration shown. The speeds v_1 and v_2 are 2.95 m/s and 5.35 m/s, respectively. The pressure P_2 is 7.36×10^4 Pa. What is P_1? (Hint: the numbers on the pressure dials are not correct - that would be too easy!)

__________ Pa



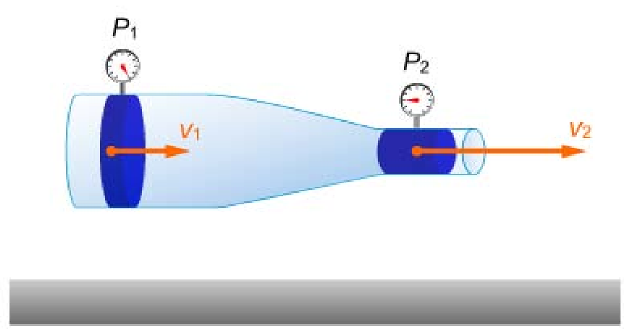
20.2 A stream of alcohol (density 790 kg/m^3) is flowing through the pipeline shown. The speeds v_1 and v_2 are 11.7 m/s and 15.8 m/s. The gauge pressures P_1 and P_2 are 2.77×10^5 Pa and 1.15×10^5 Pa, and the height h_1 is 3.60 m. What is h_2?

__________ m



20.3 A stream of water is flowing through the horizontal configuration shown. The cross-sectional areas of the left and right pipes are A_1 = 0.762 m^2 and A_2 = 0.115 m^2, and the velocity v_1 is 0.534 m/s. The pressure P_1 is 7.21×10^4 Pa. What is P_2?

__________ Pa



20.4 An open can is completely filled with water, to a depth of 20.6 cm. A hole is punched in the can at a height of 1.7 cm from the bottom of the can. Bernoulli's equation can be used to derive the following formula for the speed of the water flowing from the hole.

$$v = \sqrt{2gh}$$

In this formula, h represents the depth of the submerged hole below the surface of the water. (a) How fast does the water initially flow out of the hole? (b) How fast does the water flow when the can is half empty?

(a) __________ m/s
(b) __________ m/s

Section 23 - The Earth's atmosphere

23.1 The relationship between pressure and altitude in the Earth's atmosphere is complicated by the fact that the density of air is not constant, but decreases with height according to an equation called the **law of atmospheres**, which we do not show here. This law assumes that the atmosphere is at a constant overall Kelvin temperature T.
Because of its nonconstant density and its ill-defined upper limit, the pressure of the atmosphere cannot be described in terms of some atmospheric depth h below the "top" of the atmosphere by a simple equation like $P = \rho gh$. Instead, the atmospheric pressure P is a function of the altitude y above the ground, as given by the following equation that can be derived as an immediate consequence of the law of atmospheres:

$$P = P_{\text{atm}}\, e^{-mgy/kT}$$

In this pressure equation, P_{atm} is atmospheric pressure at sea level $(y = 0)$, m is the average mass in kilograms of an atmospheric molecule, g is the acceleration due to gravity, and k is Boltzmann's constant $(k = R/N_A = 1.38\times10^{-23}\ \text{J/K})$.
(a) If the dry atmosphere is composed of 78.1% nitrogen molecules $(m_N = 28.0\ \text{u})$, 21.0% oxygen molecules $(m_O = 32.0\ \text{u})$, and 0.930% argon molecules $(m_A = 39.9\ \text{u})$, plus trace amounts of other substances, what is the average mass of an atmospheric molecule in atomic mass units u?
(b) Convert the average atmospheric molecular mass from atomic mass units to kilograms, using the relationship 1 u = 1.66×10^{-27} kg.
(c) Now use the pressure equation given above, together with an overall atmospheric temperature of $T = 264\ \text{K}$ (which is about –9 °C), to find the theoretical pressure of the atmosphere at 1000 m, 5000 m, and 10,000 m (a common cruising altitude for commercial jetliners).

(a) __________ u
(b) __________ kg
(c) __________ Pa at 1000 m __________ Pa at 5000 m __________ Pa at 10,000 m

Section 24 - Laminar flow

24.1 Have you ever tried to suck a really thick milkshake through a thin straw? It can be a frustrating experience. (a) You want to sip water through a soda straw that has a diameter of 6.00 mm and a length of 19.0 cm, at a rate of 5.00 cm^3/s. What pressure difference do you have to apply along the length of the straw? (Assume that you are neither raising nor lowering the water you are sipping, so that laminar flow considerations are all that apply.) (b) Now, suppose you are trying to suck peanut butter through the same straw. This will be a difficult task, and you suck as hard as is physically possible, so that the pressure difference is equal to atmospheric pressure. What is the volume flow rate of the peanut butter? The coefficient of viscosity of this particular brand of peanut butter is η = 198 Pa·s.

(a) __________ Pa
(b) __________ m^3 /s

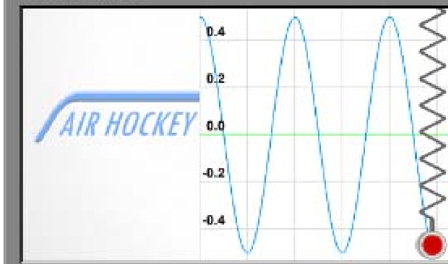
24.2 A paramedic treats an accident victim lying on the ground, transfusing her with a saline solution that has a coefficient of viscosity of $\eta = 1.002\ \text{mPa·s}$, and a density of $1025\ \text{kg/m}^3$. The flexible plastic pouch containing the saline solution is 1.85 m above the patient's arm, the needle through which the fluid is being transfused has an inside radius of 0.15 mm and a length of 1.90 cm, and the average gauge pressure of the blood in the patient's vein is $1.33\times10^4\ \text{Pa}$. (a) What is the average volume flow rate of the saline solution into the patient? (b) If this flow rate is too low, what is the first adjustment a paramedic is likely to make in the transfusion apparatus?

(a) __________ m^3 /s
(b) ○ Replace the needle with a shorter one
○ Use a needle with a larger diameter
○ Raise the height of the saline bag

chapter 15 Oscillations and Harmonic Motion

15.0 - Introduction

interactive 1



Experiment with simple harmonic motion

This chapter will give you a new take on the saying, “What goes around comes around.”

An oscillation is a motion that repeats itself. There are a myriad of examples of oscillations: a child playing on a swing, the motion of the Earth in an earthquake, a car bouncing up and down on its shock absorber, the rapid vibration of a tuning fork, the diaphragm of a loudspeaker, a quartz in a digital watch, the amount of electric current flowing in certain electric circuits, etc.!

Motion that repeats itself at regular intervals is called periodic motion. A traditional metronome provides an excellent example of periodic motion: Its periodic nature is used by musicians for timing purposes. Simple harmonic motion (SHM) describes a specific type of periodic motion. SHM provides an essential starting point for analyzing many types of motion you often see, such as the ones mentioned above.

SHM has several interesting properties. For instance, the time it takes for an object to return to an endpoint in its motion is independent of how far the object moves. Galileo Galilei is said to have noted this phenomenon during an apparently less-than-engrossing church service. He sat in the church, watching a chandelier swing back and forth during the service, and noticed that the distance the chandelier moved in its oscillations decreased over time as friction and air resistance took their toll. According to the story, he timed its period − how long it took to complete a cycle of motion − using his pulse. To his surprise, the period remained constant even as the chandelier moved less and less. (Although this is a well-known anecdote, apparently the chandelier was actually installed too late for the story to be true.)

To begin your study of simple harmonic motion, you can try the simulation to the right. A mass (an air hockey puck) is attached to a spring, and glides without friction or air resistance over an air hockey table, which you are viewing from overhead. When the puck is pushed or pulled from its rest position and released, it will oscillate in simple harmonic motion.

A pen is attached to the puck, and paper underneath it scrolls to the left over time. This enables the system to produce a graph of displacement versus time. A sample graph is shown in the illustration to the right. A mass attached to a spring is a classic configuration used to explain SHM, and the graph of the mass's displacement over time is an important element in analyzing this form of motion.

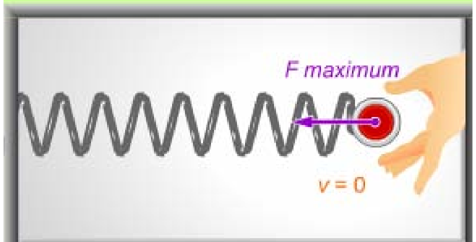
Using the controls, you can change the amplitude and period of the puck’s motion. The amplitude is the maximum displacement of the puck from its rest position. The period is the time it takes the puck to complete one full cycle of motion.

As you play with the controls, make a few observations. First, does changing the amplitude change the period, or are these quantities independent? Second, does the shape of the curve look familiar to you? To answer this question, think back to the graphs of some of the functions you studied in mathematics courses.

15.1 - Simple harmonic motion

Simple harmonic motion: Motion that follows a repetitive pattern, caused by a restoring force that is proportional to displacement from the equilibrium position.

concept 1



Simple harmonic motion
Repeated, consistent back and forth motion
Caused by a restoring force

At the right, you see an overhead view of an air hockey table with a puck attached to a spring. Friction is minimal and we ignore it. The only force we concern ourselves with is the force of the spring on the puck.

Initially, the puck is stationary and the spring is relaxed, neither stretched nor compressed. This means the puck is at its equilibrium (rest) position. Imagine that you reach out and pull the puck toward you. You see this situation in Concept 1 to the right. The spring is pulling the puck back toward its equilibrium position but the puck is stationary since you are holding onto it.

Now, you release the puck. The spring pulls on the puck until it reaches the equilibrium point. At this point, the spring exerts no force on the puck, since the spring is neither stretched nor compressed. As it reaches the equilibrium point, the puck’s speed will be at its maximum. You see this in Concept 2.

The puck’s momentum means it will continue to move to the left beyond the equilibrium point. This compresses the spring, and the force of the spring now slows the puck until it stops moving. You see this in Concept 3. At this point, the puck’s velocity is zero.

Both the displacement of the puck from the equilibrium position and the force on it are now the opposites of their starting values. The puck is as far from the equilibrium point as it was when you released it, but on the opposite side. The spring exerts an equal amount of force on the puck as it initially did, but in the opposite direction. The force will start to accelerate the puck back to the right.

The motion continues. The spring expands, pushing the puck to the equilibrium point. The puck passes this point and continues on, stretching the spring. It will return to the position from which you released it. There, the force of the stretched spring causes the puck to accelerate to the left again. Without any friction or air resistance, the puck would oscillate back and forth forever.

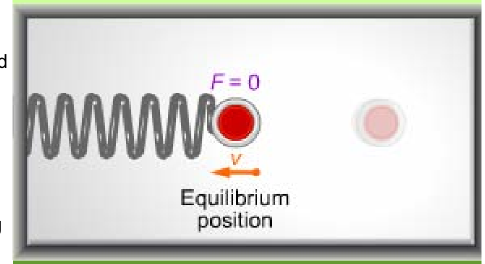
As you may have noted, “equilibrium” means there is no net force present. It does not mean “at rest” since the puck is moving as it passes through the equilibrium position. It is where the spring is neither stretched nor compressed.

The motion of the puck is called simple harmonic motion (SHM). The force of the spring plays an essential role in this motion. Two aspects of this force are required for SHM to occur. First, the spring exerts a *restoring force*. This force always points toward the equilibrium point, opposing any displacement of the puck. This is shown in the diagrams to the right: The force vectors point toward the equilibrium position.

Second, for SHM to occur, the amount of the restoring force must increase linearly with the puck’s displacement from the equilibrium point. Why can a spring cause SHM? Springs obey Hooke’s law, which states that $F = -kx$. The factor k is the spring constant and it does not vary for a given spring. As x (the displacement from equilibrium) increases in absolute value, so does the force. For instance, as the puck moves from $x = 0.25$ m to $x = 0.50$ m, the amount of force doubles. In sum, since a spring causes a restoring force that increases linearly with displacement, it can cause SHM.

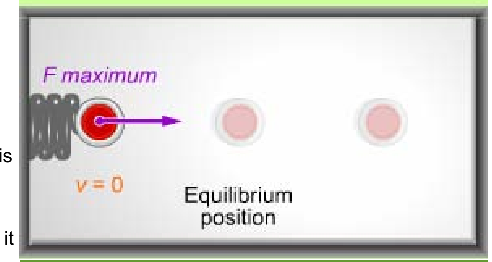
We have extensively used the example of a puck on an air hockey table here, but this is just one configuration that generates SHM. For example, we could also hang the puck from a vertical spring and allow the puck’s weight to stretch the spring until an equilibrium position was reached. If the puck were then pulled down from this position, it would oscillate in SHM, since the net force on the puck would be proportional to its displacement from equilibrium but opposite in sign.

concept 2



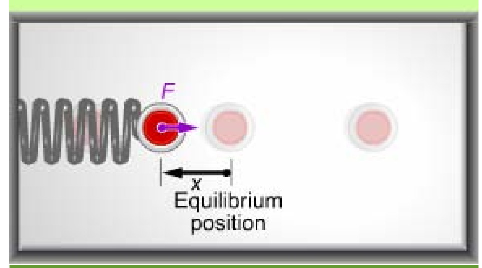
At equilibrium
Force is zero
Speed is at maximum

concept 3



Far position
Force is equal/opposite initial force
Speed is zero

concept 4



Restoring force
Proportional to displacement from equilibrium
Opposite in direction

15.2 - Simple harmonic motion: graph and equation

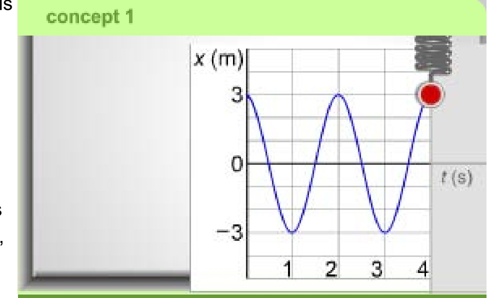
At the right, the puck is again moving in SHM, and a graph of its motion is shown. In this case, we have changed our view of the air hockey table so the puck moves vertically instead of horizontally. This puts the graph in the usual orientation. We continue to measure the displacement of the puck with the variable x, which is plotted on the vertical axis. The horizontal axis is the time t.

Unrolling the graph paper underneath the puck as it moves up and down would create the graph you see, the blue line on the white paper. The graph traces out the displacement from equilibrium of the puck over time as it moves from “peaks” where its displacement is most positive, to “troughs” where it is most negative. It starts at a peak, passes through equilibrium, moves to a trough, and so on. After four seconds, it has returned to its initial position for the second time.

The graph might look familiar to you. If you have correctly recognized the graph of a cosine function, congratulations! A cosine function describes the displacement of the puck over time. You see this function in Equation 1.

This graph represents the puck starting at its maximum displacement. When $t = 0$ seconds, the argument of the cosine function is zero radians and the cosine is one, its maximum value. (In describing SHM, the units of the argument of the cosine must be

concept 1

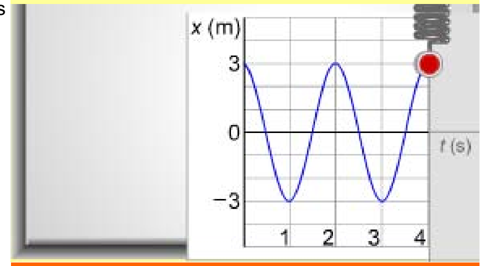


Graphing simple harmonic motion
Cosine (or sine) function describes displacement

specified as radians.) Because the function used for this graph multiplies the cosine function by an amplitude of three meters, the maximum displacement of the puck (this is always measured from equilibrium) is also three meters.

Equation 2 shows the general form of the equation for SHM. The parameters A, ω, and φ are called the *amplitude*, *angular frequency*, and *phase constant*, respectively. The argument of the cosine function, $\omega t + \varphi$, is called the *phase*. In sections that follow we will explain how these parameters are used to describe SHM.

equation 1

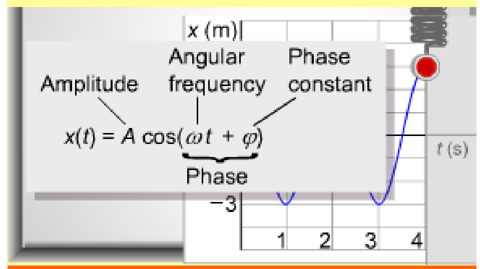


Graphing simple harmonic motion

For *this* graph:

$x(t) = 3 \cos(\pi t)$ meters

equation 2



Simple harmonic motion equation

$x(t) = A \cos(\omega t + \varphi)$

15.3 - Period and frequency

Period: Time to complete one full cycle of motion.

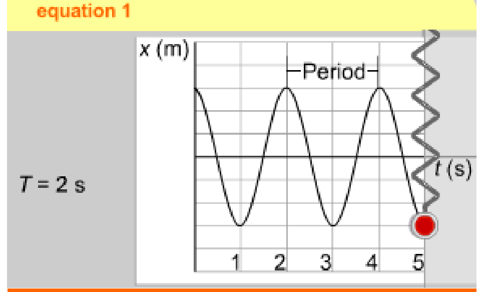
Frequency: Number of cycles of motion per second.

The period specifies how long it takes an object to complete one full cycle of motion. The letter T represents period, which is measured in seconds. A convenient way to calculate the period is to measure the time interval between two adjacent peaks, as we do in Equation 1. In the example shown there, it takes two seconds for the puck to complete one full cycle of motion.

The frequency, represented by f, specifies how many cycles are completed each second. It is the reciprocal of the period. The graph in Equation 2 is the same as in Equation 1. Its frequency is 0.5 cycles per second.

Frequency has its own units. One cycle per second equals one hertz (Hz). This unit is named after the German physicist Heinrich Hertz (1857−1894). You may be familiar with the hertz units from computer terminology: The speed of computer microprocessors used to be specified in megahertz (one million internal clock cycles per second) but microprocessors now operate at over one gigahertz (one billion cycles per second). Radio stations are also known by their frequencies. If you tune into an AM station shown on the dial at 950, the frequency of the radio waves transmitted by the station is 950 kHz.

equation 1



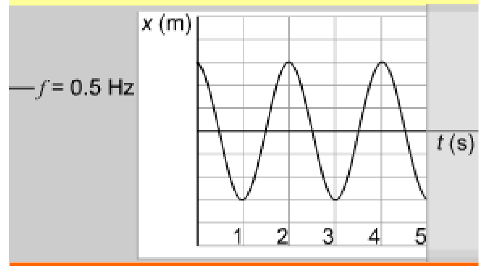
Period

Time of one complete cycle of motion

T represents period

Units: seconds (s)

equation 2

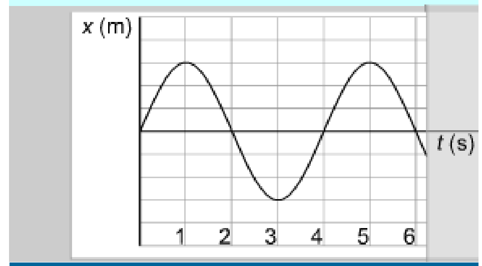


Frequency

$f = 1/T$

Cycles per second
Units: hertz (Hz)

example 1



What is the period? What is the frequency?

$T = 4.0$ s

$f = 1/T = 0.25$ Hz

15.4 - Angular frequency

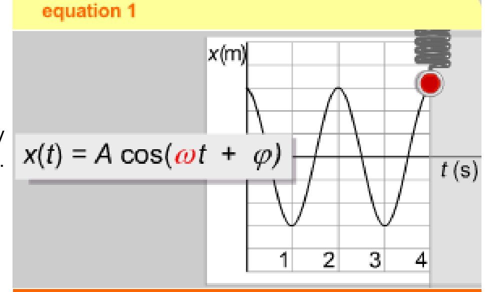
Angular frequency: Frequency measured in radians/second.

In the equation for SHM shown in Equation 1, the parameter ω is the angular frequency and it is the coefficient of time in the equation for SHM. Its units are radians per second.

The angular frequency equals 2π times the frequency. The relationship between frequency and period can be used to restate this equation in terms of the period. Both these equations are shown in Equation 2.

You may have noticed that ω also stands for the angular speed of an object moving in a circle, which is measured in radians per second, as well. If an object makes a complete loop around a circle in one second, its angular speed will be 2π radians per second. Similarly, an object in SHM that completes a cycle of motion in one second has an angular frequency of 2π radians per second. This is indicative of a relationship between circular motion and SHM that can be productively explored elsewhere.

equation 1



Angular frequency

$x(t) = A \cos(\omega t + \varphi)$

Angular frequency is ω

Units: radians per second (rad/s)

equation 2

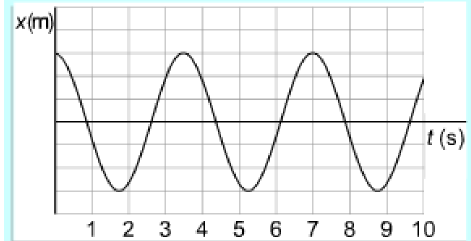
Angular frequency and period

$\omega = 2\pi f$

$\omega = 2\pi/T$

T = period
f = frequency

example 1



The function $x(t) = A\cos(\omega t)$ describes this graph. What is the angular frequency, ω?

$\omega = 2\pi/T$

T = 3.5 seconds

$\omega = 2\pi/3.5$

ω = 1.8 rad/s

15.5 - Amplitude

Amplitude: Maximum displacement from equilibrium.

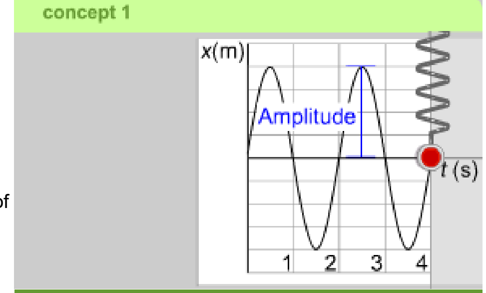
The amplitude describes the greatest displacement of an object in simple harmonic motion from its equilibrium position.

In Concept 1, you see the now familiar air hockey puck and spring, as well as a graph of its motion. The amplitude is indicated. It is the farthest distance of the puck from the equilibrium point.

The equation for SHM is shown again in Equation 1, with the amplitude term highlighted. The amplitude is the absolute value of the coefficient of the cosine function. The letter A stands for amplitude. Since the amplitude represents a displacement, it is measured in meters.

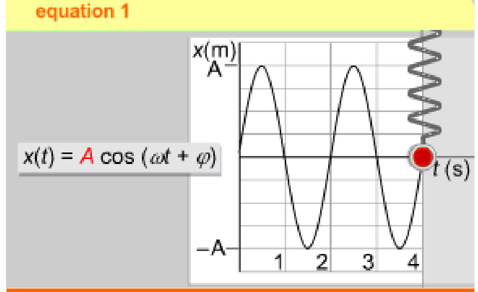
Why does the amplitude equal the factor outside the cosine function? The values of the cosine range from +1 to −1. Multiplying the maximum value of the cosine by the amplitude (for example, four meters for the function shown in Example 1) yields the maximum displacement.

concept 1



Amplitude
Maximum displacement from equilibrium

equation 1



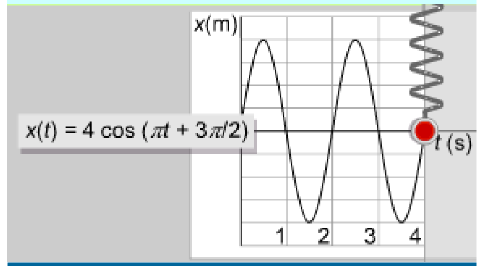
Amplitude

$x(t) = A\cos(\omega t + \varphi)$

Amplitude is $|A|$

Units: meters (m)

example 1



What is the amplitude?

Amplitude = $|A|$ = 4 m

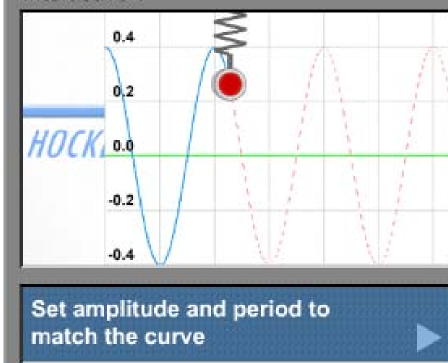
15.6 - Interactive problem: match the curve

In the simulation on the right, you control the amplitude and period for a puck on a spring moving in simple harmonic motion. With the right settings, the motion of the puck will create a graph that matches the one shown on the paper.

Determine what the values for the amplitude and period should be by examining the graph. Assume you can read the graph to the nearest 0.1 m of displacement and the nearest 0.1 s of time, and set the values accordingly. Press GO to start the action and see if your motion matches the target graph. If it does not, press RESET to try again.

Review the sections on amplitude and period if you have difficulty solving this problem.

interactive 1

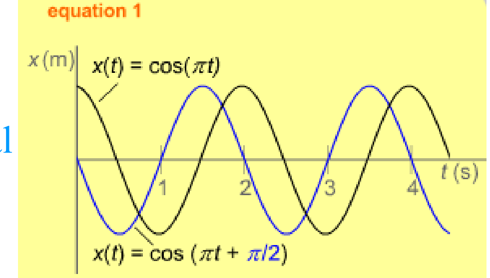


15.7 - Phase and phase constant

Phase: Argument of the trigonometric function.

Phase constant: The term of the phase that determines the value of the function at the initial time.

equation 1



Phase and phase constant

$x(t) = A\cos(\omega t + \varphi)$

Phase is $\omega t + \varphi$

Phase constant is φ

Units for φ: radians

In the general equation for SHM shown in Equation 1, the argument of the cosine function, $\omega t + \varphi$, is called the phase of the function.

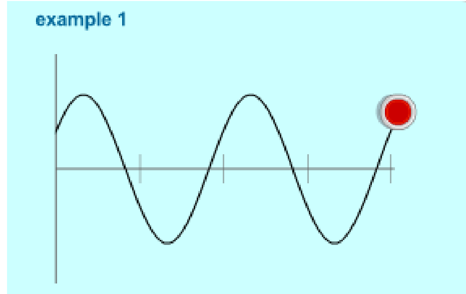
The constant term φ is called the phase constant. The phase constant can be determined by noting the displacement when $t = 0$ s and setting the phase constant appropriately given the function's amplitude. For instance, if the amplitude is 2 meters and the displacement at $t = 0$ s is 1 meter, then the phase constant must be $\pi/3$ radians $(60°)$, since the cosine of this angle is 1/2. To check this, note that $(2\text{ m})\cos(0+\pi/3) = 1\text{ m}$.

Often, it is convenient to say the object starts at its maximum positive displacement A at $t = 0$ s so that $\varphi = 0$ rad. In this chapter, we often choose examples where the phase constant is 0, to simplify the mathematical expressions.

If φ is not zero, the graph is shifted to the left or right without changing its shape. For example, if we wanted the function $x(t)$ to start at zero displacement (that is, at equilibrium), we could use a phase constant of $\pi/2$ radians $(90°)$. You see this illustrated in Equation 1. Because of the periodic nature of the cosine, you can select a phase constant in the range 0 to 2π rad to achieve any desired shift of the function. For instance, there is no difference between a phase constant of 0.5π radians, and one of 2.5π radians.

Given a graph of a simple harmonic motion function, you can calculate the phase constant. You see an example of this in Example 1. There will be two phase constants between 0 and 2π radians that result in the same displacement at $t = 0$ s, because of the identity $\cos(a) = \cos(2\pi - a)$. How can you choose between them? In the example, whether the phase constant equaled 1.1 or 5.2 radians, the graph would start at the correct initial value. To select which one is correct, you have to note whether the graph rises or falls after $t = 0$ s.

The cosine function decreases between 0 and π radians, and increases between π and 2π radians. Since the function shown in Example 1 is increasing at time $t = 0$ s, the phase constant is between π and 2π, so 5.2 is the correct choice.



An object moving in SHM has amplitude 6.3 m. At time t = 0 s, its displacement is 2.9 m. What is the phase constant, φ?

$x(t) = A \cos(\omega t + \varphi)$

$x(0\text{ s}) = 2.9\text{ m}$

$2.9\text{ m} = (6.3\text{ m}) \cos(\omega(0\text{ s}) + \varphi)$

$\cos\varphi = (2.9\text{ m})/(6.3\text{ m})$

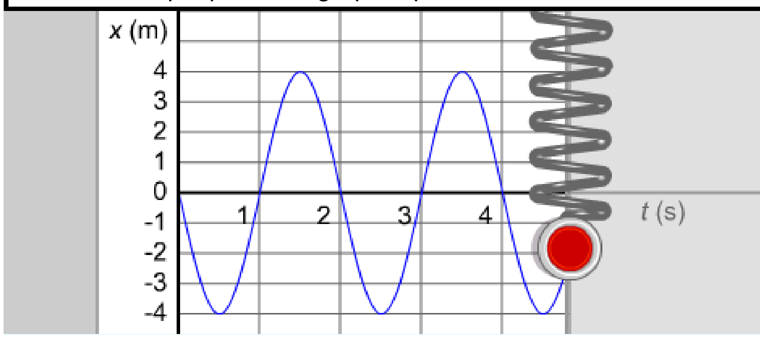
$\cos\varphi = 0.46$

$\varphi = 1.1$ rad or 5.2 rad

Because graph increases at $t = 0$,

$\varphi = 5.2$ rad

15.8 - Sample problem: graph equation



You see a graph of the displacement of an object moving in SHM. Determine the amplitude, period and phase constant, and use that information to write an equation for the displacement of the object as a function of time.

Variables

amplitude	A
period	T
angular frequency	ω
displacement	x
time	t
phase constant	φ

What is the strategy?

1. Determine the amplitude A and period T by examining the graph. From the period, calculate the angular frequency ω.
2. Examine the graph to determine the displacement at time $t = 0$, and use this and the graph to calculate the phase constant.
3. Use these values to write the SHM equation for the displacement of the object as a function of time.

Physics principles and equations

The period and angular frequency are related by

$$\omega = 2\pi/T$$

The equation for the displacement of an object in SHM is

$$x(t) = A\cos(\omega t + \varphi)$$

Step-by-step solution

The amplitude can be read directly off the graph. To determine the period, we examine the two peaks shown.

Step	Reason
1. $A = 4.0$ m	from graph
2. $T = 2.0$ s	from graph
3. $\omega = 2\pi/T$	equation for angular frequency
4. $\omega = 2\pi/2.0 = \pi$	enter value from step 2

Now we examine the graph to determine the displacement at time $t = 0$ s, and use that to calculate the phase constant. We will need to select a phase constant so that the graph decreases after $t = 0$ s.

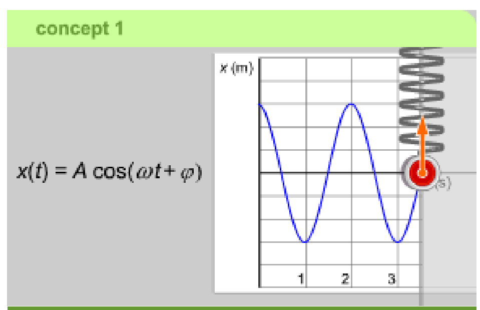
Step	Reason
5. $x(0\text{ s}) = 0.0$ m	from graph
6. $x(t) = A\cos(\omega t + \varphi)$	SHM displacement equation
7. $0.0\text{ m} = (4.0\text{ m})\cos\varphi$	from steps 1, 5, and 6
8. $\cos\varphi = 0$	solve equation 7
9. $\varphi = \pi/2$ or $3\pi/2$ rad	two possible solutions
10. $\varphi = \pi/2$ rad	graph decreasing at $t = 0$

Finally, we put the pieces together to write the equation.

Step	Reason
11. $x(t) = A\cos(\omega t + \varphi)$	SHM displacement equation
12. $x(t) = 4.0\cos(\pi t + \pi/2)$	enter values

15.9 - Velocity

In Concept 1, you see a graph of an object in simple harmonic motion. The graph shows the displacement of the object versus time. At any point the slope of the graph is the object's instantaneous velocity. The slope equals $\Delta y/\Delta x$. In this graph, this is the change in displacement per unit time, which is velocity.



Velocity in SHM
Velocity constantly changes
· Extreme velocities at equilibrium
· Zero velocity at endpoints

You can consider the relationship of velocity and displacement by reviewing the role of force in SHM. Consider an object attached to a spring, where the spring is stretched and then the object is released. The spring force pulls the object until it reaches the equilibrium point, increasing the object's speed.

Once the object passes through the equilibrium point, the spring is compressed and its force resists the object's motion, slowing it down. Because the object speeds up as it approaches the equilibrium point and slows down as it moves away from equilibrium, its greatest speed is at the equilibrium point.

When the spring reaches its maximum compression, the object stops for an instant. At this point, its speed equals zero. The spring then expands until the object returns to its initial position, with the spring fully extended. Again, the object stops for an instant, and its speed is zero.

In the paragraphs above, we discussed the motion in terms of speed, not velocity, so we could ignore the sign and focus on how fast the object moves. The object's velocity will be both positive and negative as it moves back and forth. You see this alternating pattern of positive and negative velocities in the graph in Equation 1.

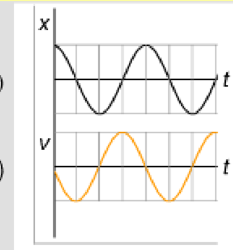
When the displacement is at an extreme, the velocity is zero, and vice-versa. One way to state the relationship between the displacement and velocity functions is to say they are $\pi/2$ radians $(90°)$ out of phase. An equivalent way to express this without a phase constant is to use a cosine function for displacement and a sine function for velocity, and this is what we do. This relationship can also be derived using calculus. In Equation 1, you see both a velocity graph and the function for velocity.

The second equation shown in Equation 1 states that the maximum speed v_{max} is the amplitude of the displacement function times the angular frequency. To understand the source of this equation, recall that the maximum magnitude of the sine function is one. When the sine has a value of −1 in the velocity equation, the velocity reaches its maximum value of $A\omega$.

equation 1

$x(t) = A\cos(\omega t + \varphi)$

$v(t) = -A\omega \sin(\omega t + \varphi)$



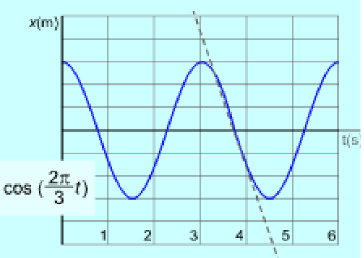
Velocity in SHM

$$v(t) = -A\omega \sin(\omega t + \varphi)$$
$$v_{max} = A\omega$$

v = velocity
A = amplitude
ω = angular frequency
t = time
φ = phase constant

example 1

$x(t) = 3.0 \cos\left(\frac{2\pi}{3}t\right)$



What is the velocity at t = 4.0 seconds?

$v(t) = -A\omega \sin(\omega t + \varphi)$

$$v(4.0\ \text{s}) = (-3.0\ \text{m})\left(\frac{2\pi}{3}\frac{\text{rad}}{\text{s}}\right)\sin\left(\left(\frac{2\pi}{3}\frac{\text{rad}}{\text{s}}\right)(4.0\ \text{s})\right)$$

$v(4.0\ \text{s}) = (-2\pi) \sin(8\pi/3)$ m/s

$v(4.0\ \text{s}) = (-6.28)(0.866)$ m/s

$v(4.0\ \text{s}) = -5.4$ m/s

15.10 - Interactive checkpoint: particle speed

A particle vibrates in simple harmonic motion with a frequency of 2.60×10^{9} Hz and an amplitude of 1.70×10^{-8} m. What is its maximum speed?

Answer:

v_{max} = ______ m/s

15.11 - Acceleration

For SHM to occur, the net force on an object has to be proportional and opposite in sign to its displacement. Again, we use the example of a mass attached to a spring on a friction-free surface, like an air hockey table.

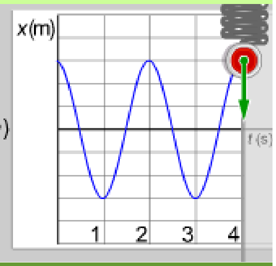
With a spring like the one shown in Concept 1 to the right, Hooke's law $(F = -kx)$ states the relationship between net force and displacement from equilibrium. This equation for force enables you to determine where the acceleration is the greatest, and where it equals zero. The magnitudes of the force and the acceleration are greatest at the extremes of the motion, where x itself is the greatest. This is the point where the object is changing direction. Conversely, $x = 0$ at the equilibrium point, so $F = 0$ and the object is not accelerating there.

The first equation shown in Equation 1 enables you to calculate the acceleration of an object in SHM as a function of time. This equation can be simplified by noting that the amplitude times the cosine function, the rightmost term in the equation, is the function for the object's displacement, $x(t)$. We replace the terms $A\cos\omega t$ by $x(t)$ to derive the second equation, which relates the acceleration directly to the object's displacement. This equation says that the acceleration at a particular time equals the negative of the angular frequency squared times the object's displacement at that time.

Finally, the third equation reveals that the maximum acceleration of the object is the amplitude times the square of the angular frequency. This equation is a consequence of the first equation.

concept 1

$x(t) = A\cos(\omega t + \varphi)$



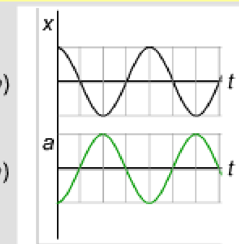
Acceleration in SHM

Proportional to force
· Zero at equilibrium
· Maximum at extremes

equation 1

$x(t) = A\cos(\omega t + \varphi)$

$a(t) = -A\omega^2\cos(\omega t + \varphi)$



Acceleration in SHM

$a(t) = -A\omega^2\cos(\omega t + \varphi)$

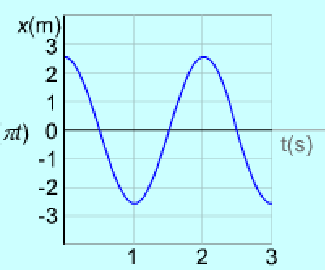
$a(t) = -\omega^2 x(t)$

$a_{\text{max}} = A\omega^2$

a = acceleration, A = amplitude
ω = angular velocity
x = displacement, t = time
φ = phase constant

example 1

$x(t) = 2.5\cos(\pi t)$



What is the acceleration at t = 1.6 seconds?

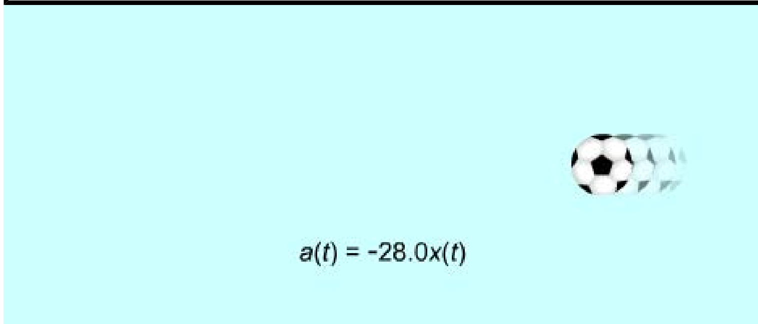
$a(t) = -A\omega^2\cos(\omega t + \varphi)$

$a(1.6\text{s}) =$

$-2.5\text{m}\left(\pi\frac{\text{rad}}{\text{s}}\right)^2\cos\left(\pi\frac{\text{rad}}{\text{s}}\cdot 1.6\text{s}\right)$

$a(1.6\text{ s}) = -7.6\text{ m/s}^2$

15.12 - Sample problem: calculating period from acceleration



The ball moves in SHM, and its acceleration is defined by the equation shown. What is the period of the ball's motion?

Variables

acceleration	$a(t) = -28.0\,x(t)$
angular frequency	ω
period	T

What is the strategy?

1. Use the equation relating acceleration to angular frequency and displacement, and the given equation, to calculate the angular frequency.
2. Calculate the period from the angular frequency.

Physics principles and equations

The acceleration of an object in SHM is given by

$$a(t) = -\omega^2 x(t)$$

Period and angular frequency are related by the equation

$$\omega = 2\pi/T$$

Step-by-step solution

There are two equations for acceleration: the one given in the problem, and one that relates acceleration to angular frequency and displacement. We use these equations to calculate the angular frequency.

Step	Reason
1. $a(t) = -\omega^2 x(t)$	acceleration equation
2. $a(t) = -(28.0)x(t)$	equation stated in problem
3. $-\omega^2 x(t) = -(28.0)x(t)$	set right sides of steps 1, 2 equal
4. $\omega^2 = 28.0$	divide by $-x(t)$
5. $\omega = 5.29$ rad/s	square root

Now we calculate the period from the angular acceleration.

Step	Reason
6. $\omega = 2\pi/T$	equation for period
7. $T = 2\pi/\omega$	solve for T
8. $T = 2\pi/(5.29\text{ rad/s})$	enter angular acceleration
9. $T = 1.19$ s	evaluate

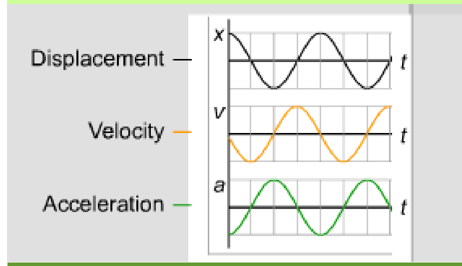
15.13 - Summary of simple harmonic motion

You have now been exposed (subjected?) to three trigonometric functions, and three graphs, for the displacement, velocity and acceleration of an object in simple harmonic motion. All of these equations use the sine or cosine function; these are called *sinusoidal functions*. We show the equations for these properties of motion in Equation 1.

You may find it useful to see the three graphs of these functions together, as shown in Concept 1. Remember that the vertical axes represent different units. Going from the top graph to the bottom one, the axes are meters, m/s, and m/s^2.

The graphs illustrate the relationship of displacement, velocity and acceleration. You can "see" some of the nature of simple harmonic motion in these graphs. For example, when the particle is at one of its endpoints, displacement has the greatest magnitude, and velocity is zero as the particle momentarily pauses and changes the sign of its velocity. When the displacement has its greatest positive value, the acceleration has its most negative value, and vice versa.

concept 1



Summary of SHM
Sinusoidal functions describe displacement, velocity, acceleration

equation 1

Summary of SHM

$$x(t) = A \cos(\omega t + \varphi)$$
$$v(t) = -A\omega \sin(\omega t + \varphi)$$
$$a(t) = -A\omega^2 \cos(\omega t + \varphi)$$

x = displacement, v = velocity
a = acceleration, A = amplitude
ω = angular frequency
t = time, φ = phase constant

15.14 - Simple harmonic motion and uniform circular motion

Galileo was the first to make observations of the moons of Jupiter. He used a telescope, at that time a recent invention, to discover four of the planet's moons. He noted that the moons move in a pattern that could be best explained if the objects he observed were moving in circles around Jupiter. Using today's vocabulary, we would say that the pattern of motion he observed could be described using the concepts of simple harmonic motion. This is not to say the moons move in SHM, because they do not. But they appear to be doing so because an object moving in uniform circular motion viewed edge-on from afar appears to be moving in SHM.

To help you understand Galileo's insight, we will explain in more depth what he observed. The moons' orbits are roughly circular. From his perspective on Earth, Galileo could only see the moons' lateral motion, their motion from left to right and right to left. At the moons' great distance, their motion toward and away from the Earth was not perceptible. The moons seem to repeatedly move back and forth along a straight line, as objects in SHM do.

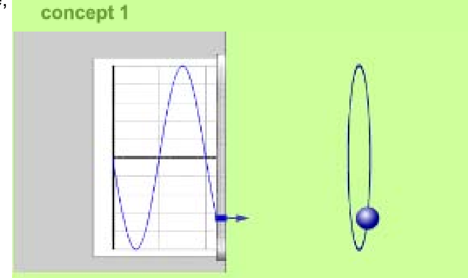
When an object moving in uniform circular motion, such as a moon of Jupiter, is observed "edge on," does the repeated motion actually conform to the equations for SHM? In Concept 1 we show a ball moving up and down in circular motion. The graph shows its vertical displacement and, as you can see, the graph is sinusoidal, like those that represent displacement in SHM.

In Equation 1, you can see why we perceive edge-on uniform circular motion as SHM. Consider the x displacement of the particle moving in uniform circular motion. The x displacement equals the radius (which we will call A here) times the cosine of the angle θ. The angle θ is the angular displacement. As you may recall from your studies of rotational motion, angular displacement equals the product of angular velocity and time, or ωt. As the equation to the right reflects, the function for the x displacement is the same function as the one used for calculating the displacement of an object in SHM.

The relationship between uniform circular motion and SHM can also be confirmed qualitatively. At the perceived endpoints of its circular path, a moon of Jupiter would be moving directly away from, or toward, Galileo. In other words, its tangential velocity is directed toward or away from the Earth at these points. Here its velocity should seem to go to zero just as in SHM, and this is what Galileo observed.

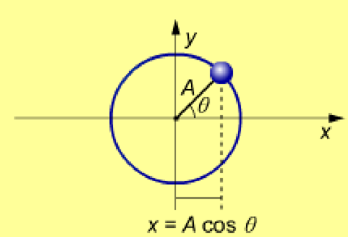
Conversely, at the midpoint of its motion on the Earth side of Jupiter, a moon would seem to be moving the fastest because of the orientation of its tangential velocity. And

concept 1



SHM and uniform circular motion
Uniform circular motion viewed edge-on appears as SHM

equation 1



SHM and uniform circular motion

$$x(t) = A \cos\theta = A \cos\omega t$$

θ = angular displacement

in SHM, the greatest speed is observed at the midpoint of motion.

ω = angular velocity
t= time

15.15 - Period, spring constant, and mass

On Skylab, astronaut Alan Bean used the period of a spring-driven chair to measure his mass.

In Equation 1, you see two equations that relate the physical properties of a mass-spring system moving in SHM to its angular frequency and period. These reflect how the physical properties of this system determine its motion. These equations are applicable when the mass of the spring is negligible compared to the object attached to it.

The first equation states that the angular frequency equals the square root of the spring constant divided by the mass. You can think of this equation in terms of Hooke's law and Newton's second law. Hooke's law states that for a given displacement, the greater the spring constant, the greater the force. This means there will be greater acceleration and the mass will both reach its peak velocity and return to zero velocity more quickly. This gives you reason to think the frequency of motion increases with the spring constant.

In contrast, by Newton's second law acceleration is inversely proportional to mass. As the mass increases the acceleration decreases, so angular frequency decreases with mass.

The second equation, an equation for period, can be derived from the first equation using the relationship $\omega = 2\pi/T$. Since the period is inversely proportional to the angular frequency, the period increases as mass increases and decreases as the spring constant increases.

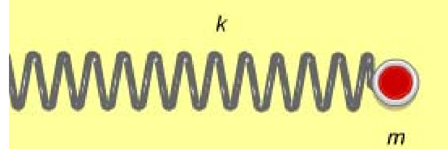
Notice that the angular frequency and the period depend only on the mass and spring constant, not on the amplitude. When the amplitude increases, the mass has to travel farther during a cycle. However, increasing the amplitude also increases the maximum force applied to the mass, which increases its maximum acceleration and results in a greater average speed. The increased average speed means the mass completes the cycle in the same amount of time.

To derive the first equation, we use Newton's second law and Hooke's law. We also use an equation discussed earlier to calculate the acceleration as a function of the angular frequency. Then using algebra we can solve for the angular frequency.

Step	Reason
1. $F = ma$	Newton's second law
2. $a = -\omega^2 x$	SHM acceleration
3. $F = -m\omega^2 x$	substitute into equation 1
4. $F = -kx$	Hooke's law
5. $-m\omega^2 x = -kx$	from steps 3 and 4
6. $\omega = \sqrt{k/m}$	solve for ω

In Example 1, we calculate the period of a mass on a spring.

equation 1



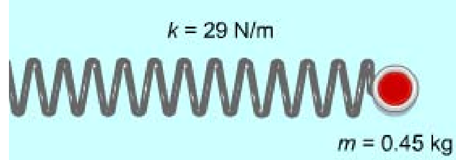
Angular frequency and period

$$\omega = \sqrt{\frac{k}{m}}$$

$$T = 2\pi\sqrt{\frac{m}{k}}$$

ω = angular frequency
T = period
k = spring constant
m = mass

example 1



What is the period of the puck's motion?

$$T = 2\pi\sqrt{\frac{m}{k}}$$

$$T = 2\pi\sqrt{\frac{0.45\text{ kg}}{29\text{ N/m}}}$$

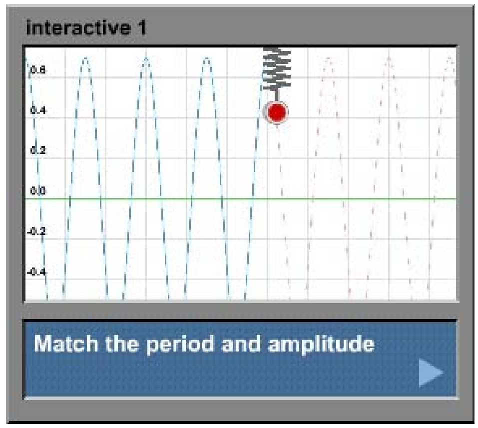
$$T = 0.78\text{ s}$$

15.16 - Interactive problem: match the curve again

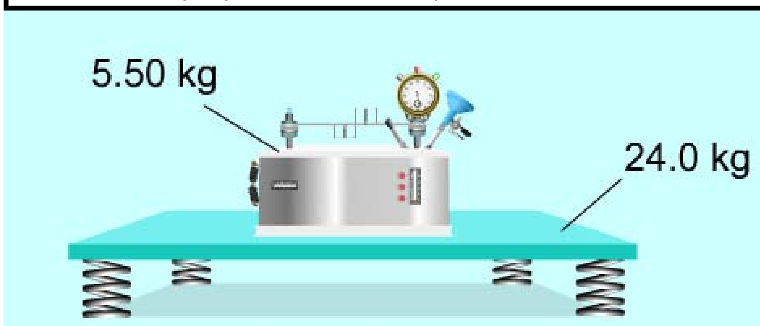
In the simulation on the right, you control parameters for a puck on a spring moving in simple harmonic motion. With the right settings, the puck will create a graph that matches the one shown on the paper.

You can determine the amplitude by examining the graph. The spring constant is 10.0 N/m. The desired period can be determined by examining the graph, but you do not set the period directly. Instead, you set the mass of the puck, which determines the period. Use the spring constant and the period to calculate what the mass should be. Enter the amplitude to the nearest 0.01 m and the mass to the nearest 0.01 kg and press GO. A gauge shows the actual period so that you can confirm your calculations. If your graph does not match the target graph, press RESET to try again.

You should review the section on springs and period if you have trouble solving this problem.



15.17 - Sample problem: isolation platform



The isolation platform is supported by four springs, each with spring constant 425 N/m. What is the natural frequency of vibration of the platform and equipment?

Laboratories isolate sensitive equipment from floor vibrations by using a platform resting on springs, as you see above. The springs are chosen so that the frequency of vibration of the platform plus the equipment is below 2 Hz. The frequency of vibrations from the floor is typically in the 5−30 Hz range. This difference in frequencies allows the isolation platform to "filter out" the floor vibrations before they reach the equipment. In the platform above, each of the four identical springs supports one-fourth of the total mass.

Variables

mass of platform	$m_1 = 24.0$ kg
mass of equipment	$m_2 = 5.50$ kg
spring constant for a spring	$k = 425$ N/m
angular frequency of system	ω
frequency of a spring	f

What is the strategy?

Since the springs are identical, and each supports an equal fraction of the total mass (and weight), they have the same frequency. We calculate the frequency of one spring as follows.

1. Calculate the angular frequency of a system consisting of one spring, supporting one-fourth of the total mass.
2. Calculate the frequency from the angular frequency.

Physics principles and equations

The angular frequency of a mass on a spring is

$$\omega = \sqrt{\frac{k}{m}}$$

Frequency and angular frequency are related by the equation

$$\omega = 2\pi f$$

Step-by-step solution

We start by considering a system consisting of one spring supporting one-fourth the total mass.

Step	Reason
1. $\omega = \sqrt{\dfrac{k}{m}}$	equation for angular frequency of mass on spring
2. $\omega = \sqrt{\dfrac{k}{(m_1+m_2)/4}}$	mass supported by one spring
3. $\omega = \sqrt{\dfrac{425 \text{ N/m}}{(24.0 \text{ kg} + 5.50 \text{ kg})/4}}$	enter values
4. $\omega = 7.59$ rad/s	evaluate

Now that we have the angular frequency of the system, we can calculate the frequency.

Step	Reason
5. $\omega = 2\pi f$	relationship of angular frequency and frequency
6. $f = \omega/2\pi$	solve for f
7. $f = (7.59 \text{ rad/s})/2\pi$	enter angular frequency
8. $f = 1.21$ Hz	evaluate

The frequency of 1.21 Hz is in the desirable range for an isolation platform.

15.18 - Work and the potential energy of a spring

We will once again use a spring as our tool for analyzing simple harmonic motion. Springs can store potential energy. The energy stored by a spring is called *elastic potential energy*.

When the spring shown in Concept 1 is extended, it has elastic potential energy. When the hand releases the mass at the end of the spring, potential energy will be converted into kinetic energy as the spring and mass move.

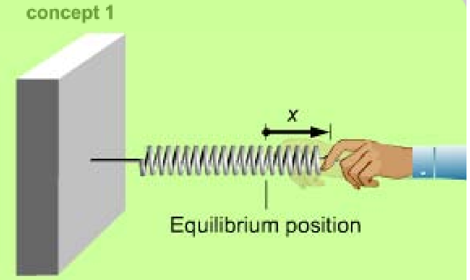
It takes work to stretch or compress a spring. The work done on the spring-mass system will equal the **change** in total energy of the system. It is useful to consider the system when it has **only** potential energy, for instance, when we pull back on the mass from its rest (equilibrium) position and hold it still. We define the potential energy of the system to be zero when the mass is at its equilibrium position, so the work we do on the system as we stretch the spring and then hold it equals the potential energy stored in the stretched spring.

The formula in Equation 1 states that the potential energy equals one-half the spring constant k times the square of the displacement x. Although we use the example of the spring and the spring constant, this expression can be used to calculate the potential energy of any system with a restoring force proportional to displacement, with the appropriate restoring force constant k.

We can graphically derive the value for potential energy shown in the equation. In Equation 1, you see a graph that shows how the displacement, force and work are related.

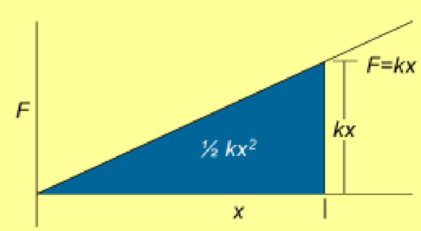
1. The potential energy stored by the spring when it is stretched equals the work done **on** the system by the pulling hand. At any displacement x, the hand exerts a force kx, opposing the spring force. The linear graph shows this force as a function of x.
2. When force varies with displacement as it does here, the work done equals the area under the graph. The region under the graph is a right triangle whose area equals one-half the product of its base times its height. The base of the triangle is the displacement x and the height is the force kx. The area of the triangle, then, is ½ kx^2, which equals the work done, and the PE as well.

concept 1



Work and potential energy
Elastic potential energy stored in system
Equals work done on system

equation 1



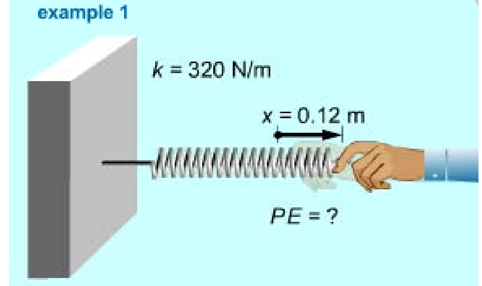
Spring potential energy

$PE = \frac{1}{2}kx^2$

PE = potential energy
k = spring constant

x = displacement from equilibrium

example 1



How much potential energy is stored in the spring?

$PE = \frac{1}{2}kx^2$

$PE = \frac{1}{2}(320\ \text{N/m})(0.12\ \text{m})^2$

$PE = 2.3\ \text{J}$

15.19 - Total energy

When an object on a spring moves in simple harmonic motion, the system's energy changes back and forth from elastic potential energy to kinetic energy. Assuming no non-conservative forces like friction are present, the law of conservation of energy dictates that the total energy of the system remains constant over time.

In Concept 1, you see a graph of the displacement of the puck as it oscillates in SHM. There are also three energy gauges for the mass/spring system. As the puck moves and the spring expands and contracts, the values displayed in the PE and KE gauges will change, but their sum, the total energy (TE), remains constant.

At its maximum displacement, the puck has zero kinetic energy because it has zero velocity. All of the energy of the system is elastic potential energy in the spring. This is the situation shown in Concept 1. At the equilibrium position, where the puck is moving the fastest and the spring is relaxed, all of the system's energy is kinetic energy. We have defined the equilibrium position as the point of zero potential energy.

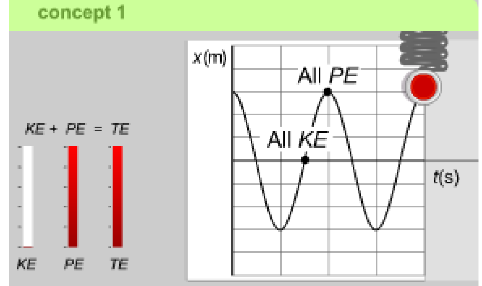
In Equation 1, you see equations for the two kinds of energy in a system consisting of a mass m and a spring with constant k. The kinetic energy equals $\frac{1}{2}mv^2$ as always, but here we express its velocity using the sinusoidal velocity expression for an object in SHM.

As was shown in the prior section, the potential energy arising from the restoring force of an object in simple harmonic motion is $\frac{1}{2}kx^2$, where x is the displacement of the object. In the second equation in Equation 1, we write this using the sinusoidal expression for the displacement of an object in SHM.

We combine the graphs for KE and PE in Equation 2. The values for the graphs sum to a constant at each point. The total energy of the system does not change, even as its components do.

To determine an equation for total energy, consider the puck when it is at its farthest point from equilibrium, where $x = A$. At this point, the puck has zero velocity, so there is no kinetic energy. The total energy is solely the potential energy, which equals $\frac{1}{2}kA^2$. As the puck begins to move, its KE increases as its PE decreases, but the total energy remains the same. We show this as an equation in Equation 2 on the right.

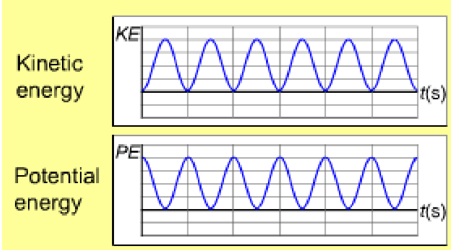
concept 1



Total energy

Total energy is constant (*KE* + *PE*)
At equilibrium: all *KE*
At maximum displacement: all *PE*

equation 1



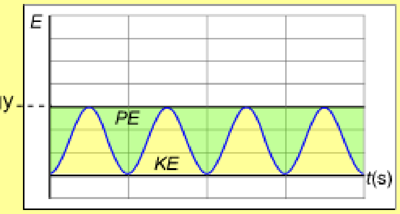
Kinetic, potential energy

$KE = \frac{1}{2}mA^2\omega^2 \sin^2 \omega t$

$PE = \frac{1}{2}kA^2 \cos^2 \omega t$

m = mass
A = amplitude
ω = angular frequency
t = time
k = restoring force constant

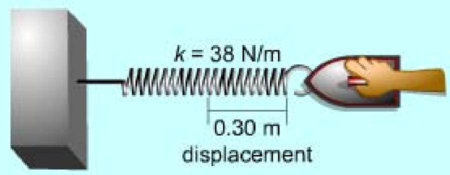
equation 2



Total energy

$TE = \frac{1}{2}kA^2$

example 1



When the friction-free iron is released, what will its *KE* be at its equilibrium point?

Total energy at equilibrium is all KE

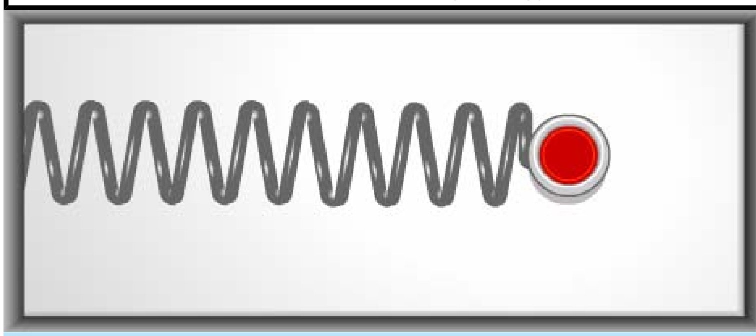
KE at equilibrium = initial PE

$KE = \frac{1}{2}kA^2$

$KE = \frac{1}{2}(38 \text{ N/m})(0.30 \text{ m})^2$

$KE = 1.7 \text{ J}$

15.20 - Interactive checkpoint: spring energy and period

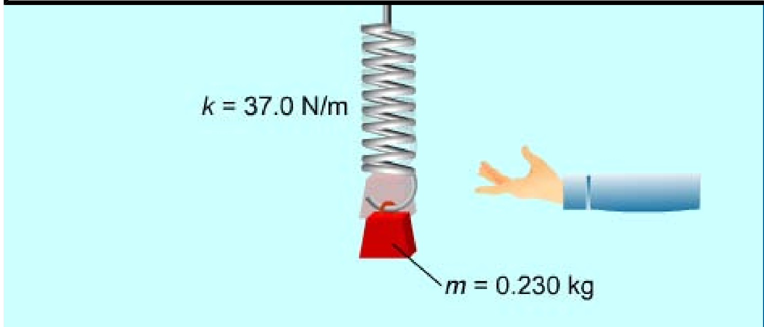


A disk mounted on a massless spring oscillates in simple harmonic motion. The disk has a mass of 0.580 kg and moves with an amplitude of 0.150 m. The total energy of the system is 1.30 J. What is the system's period?

Answer:

T = ______ seconds

15.21 - Sample problem: falling block on a spring



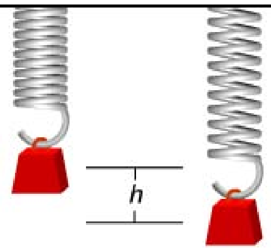
The block is held so the spring is neither stretched nor compressed. When the block is released, how far does it move before reversing direction?

In many prior examples, we showed a mass attached to a spring on a frictionless horizontal surface. The only force we needed to consider was the restoring force from the spring. Here, since the mass is attached vertically to a spring, we must account for the force of gravity as well.

The two forces acting on the block are the downward force of gravity and the upward force of the spring (we ignore other forces like air resistance). The distance the block falls before stopping can be calculated using the conservation of energy. In this case, we need to consider gravitational potential energy as well as the potential energy in the spring. We ignore the mass of the spring itself.

Draw a diagram

We let h represent the distance the block falls. We consider the gravitational potential energy of the block to be zero at the bottom of the block's motion.



Variables

mass of block	$m = 0.230$ kg
spring constant	$k = 37.0$ N/m
initial block height	h
final block height	0 m
initial total energy of block	TE_i
final total energy of block	TE_f

What is the strategy?

1. Use the conservation of energy to state that the total energy before the block is released is the same as the total energy at the turnaround point.
2. The block is not moving initially, and is (momentarily) not moving when it changes direction. Its kinetic energy is zero at both of these positions. Use this fact to simplify the equation.
3. Solve for the distance h.

Physics principles and equations

Conservation of energy

$$TE_i = TE_f$$

Total energy of an object in simple harmonic motion

$$TE = PE + KE$$

Gravitational potential energy

$$PE = mgh$$

Elastic potential energy from a spring

$$PE = ½kx^2$$

Step-by-step solution

Since the block has zero velocity at both the top and bottom positions of its motion, its initial and final kinetic energies are both zero. The total energy, in both the initial and final states, equals the potential energy. The initial potential energy of the block is all in the form of gravitational potential energy since the spring is unstretched. The potential energy as the block changes direction at the bottom of its motion is all due to the elastic potential energy in the spring, because we set the gravitational PE to zero there.

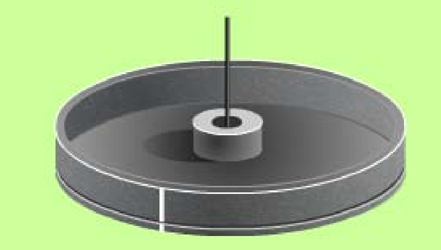
Step	Reason
1. $TE_i = TE_f$	conservation of energy
2. $PE_i + KE_i = PE_f + KE_f$	total energy of object
3. $PE_i = PE_f$	kinetic energies are zero
4. $mgh = ½kh^2$	enter expressions for potential energies

Now we have an equation where the only unknown value is the one we want to calculate.

Step	Reason
5. $h = \dfrac{2mg}{k}$	solve for h
6. $h = \dfrac{2\,(0.230\text{ kg})(9.80\text{ m/s}^2)}{37.0\text{ N/m}}$	enter values
7. $h = 0.122$ m	evaluate

15.22 - A torsional pendulum

concept 1



Torsional pendulum
Exhibits simple harmonic motion
Use rotational concepts to analyze

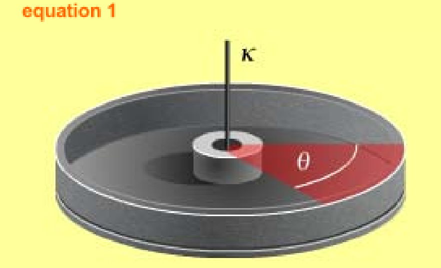
The *torsional pendulum* shown in Concept 1 is another device that exhibits simple harmonic motion. A torsional pendulum consists of a mass suspended at the end of a stiff rod, wire or spring. It does not swing back and forth. Instead, the mass at the bottom is initially rotated by an external torque away from its equilibrium position. The elasticity of the rod supplies a *restoring torque*, causing the mass to rotate back to the equilibrium position and beyond. The mass rotates in an angular version of simple harmonic motion.

Earlier, we stated that for SHM to occur, the force must be proportional to displacement. Since torsional pendulums rotate, we must use angular concepts to analyze them. With a torsional pendulum, a restoring torque, not a force, acts to return the system to its equilibrium position. The restoring torque is proportional to angular displacement, just as a restoring force is proportional to (linear) displacement. The moment of inertia of the system takes on the role that mass plays in linear SHM.

The same analysis that applies to linear displacement, velocity and acceleration applies equally well to angular displacement, angular velocity and angular acceleration. In Equation 1, you see an equation that states the nature of the restoring torque. It equals the negative of the product of the *torsion constant* and the angular displacement.

The formula in Equation 2 calculates the period of the pendulum. When the period and the torsion constant are known, the moment of inertia can be calculated, as shown in Example 1. This makes torsional pendulums useful tools for experimentally determining the moments of inertia of complex objects.

equation 1



Restoring torque

$$\tau = -\kappa\theta$$

τ = torque
κ = torsion constant
θ = angular displacement

Units for κ: N·m/rad

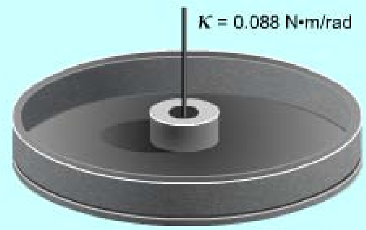
equation 2

Period

$$T = 2\pi\sqrt{\frac{I}{\kappa}}$$

T = period
I = moment of inertia
κ = torsion constant

example 1



The torsional pendulum has a period of 3.0 s. What is its moment of inertia?

$$T = 2\pi\sqrt{\frac{I}{\kappa}}$$

$$I = \frac{\kappa T^2}{4\pi^2}$$

I = (0.088 N·m/rad)(3.0 s)^2/4π^2
I = 0.020 kg·m^2

15.23 - A simple pendulum

Old-fashioned "grandfather" clocks, like the one you see in Concept 1, rely on the regular motion of their pendulums to keep time. A typical pendulum is constructed with a heavy weight called a "bob" attached to a long, thin rod. The bob swings back and forth at the end of the rod in a regular motion.

We approximate such a system as a *simple pendulum*. In a simple pendulum, the bob is assumed to be concentrated at a single point located at the very end of a cable, and the cable itself is treated as having no mass. The system is assumed to have no friction and to experience no air resistance. When such a pendulum swings back and forth with a small amplitude, its angular displacement closely approximates simple harmonic motion. This means the period does not vary much with the pendulum's amplitude. This regularity of period is what makes pendulums useful in clocks.

For SHM to occur, the restoring force or torque needs to vary linearly with displacement. In the case of a pendulum, the motion is rotational, so the torque must be linearly proportional to the angular displacement.

concept 1

Simple pendulum
Point mass at end of massless rod
Approximates simple harmonic motion

In Equation 1, you see a free-body diagram of the forces on the pendulum bob. The tension in the cable exerts no torque on the pendulum since it passes through its center of rotation, so the weight mg of the bob exerts the only torque. The lever arm of this weight equals the length L of the cable times $\sin\theta$. For small angles, the angle expressed in radians is a very close approximation of the sine of the angle. (The error is less than 1% for angles less than 14°.) This means that the resulting torque is roughly proportional to the angular displacement, and the condition for SHM is approximated, with a torsion constant of mgL.

In Equation 2, you see the equation for the period of a simple pendulum. When the angular amplitude is small and the approximation mentioned above is used, the period depends solely on the length of the cable and the acceleration of gravity.

A pendulum can be an effective tool for measuring the acceleration caused by gravity using the equation just mentioned. The length L of the cable is measured and the pendulum is set swinging with a small amplitude. The period T is then measured. The value of g can be calculated using the rearranged equation $g = 4L(\pi/T)^2$.

equation 1

Restoring torque

$$\tau = -mgL \sin\theta \approx -mgL\theta$$

τ = torque
m = mass
g = acceleration of gravity
L = length of pendulum
θ = angular displacement

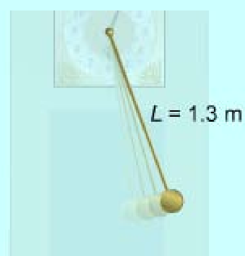
equation 2

Period

$$T = 2\pi\sqrt{\frac{L}{g}}$$

T = period
L = length of pendulum
g = acceleration of gravity

example 1



What is the period of this pendulum?

$$T = 2\pi\sqrt{\frac{L}{g}}$$

$$T = 2\pi\sqrt{\frac{1.3\text{ m}}{9.80\text{ m/s}^2}}$$

$$T = 2.3\text{ s}$$

15.24 - Interactive problem: a pendulum

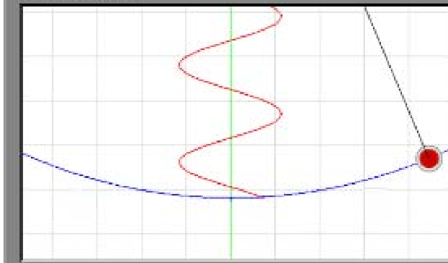
On the right is a simulation of a simple pendulum: a bob at the end of a string. You can control the length of the string, and in doing so change the period of the pendulum. Your goal is to set the length so that the period is 2.20 seconds. As the pendulum swings, you will see a graph reflecting the **angular** displacement of the bob.

Calculate and set the value for the string length to the nearest 0.05 m using the dial, then use your mouse to drag the bob to one side and release it to start the pendulum swinging. There may not be enough room in the window to show the entire length of the string, but we will show the motion of the bob and the resulting period. If you do not set the length correctly, press RESET to try again. Refer to the section on simple pendulums if you do not remember the equation for the period.

You may want to double-check your work by creating an actual pendulum with a string of the correct length. You can time it: Ten cycles of its motion should take about 22 seconds.

For small angles, the angular displacement of a pendulum approximates simple harmonic motion and the graph looks sinusoidal. Try smaller and larger angles and observe the graphs. How sinusoidal do they look to you? (In the simulation, decreasing the string length makes it easier to create large angular displacements.) You can check the box labeled "SHM" to draw a sinusoidal graph of SHM motion in black underneath your red graph. The black graph shows simple harmonic motion for the amplitude you choose and the period calculated by the pendulum equation. If the amplitude is small, you might not see the black graph, because the two graphs match so closely.

interactive 1



Set the pendulum length for a 2.20 second period

15.25 - Period of a physical pendulum

Not all pendulums are simple. A *physical pendulum* is a rigid extended object (not a point mass) pivoting around a point. In Concept 1, you see a violin acting as a physical pendulum.

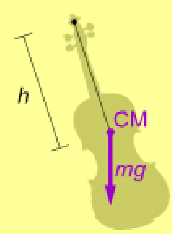
The equation for the period of a physical pendulum is shown in Equation 1. The distance h is the distance from the pivot point to the center of mass of the object. As always, the moment of inertia must be calculated about the pivot point.

The example problem shows how to calculate the period of a meter stick used as a pendulum in the Earth's gravitational field. The period is 1.6 seconds. With the use of a meter stick, this is a result you can verify for yourself. If the stick has a hole close to one end, put an unbent paper clip through the hole (otherwise, pinch the end very loosely between your fingers), and set the stick swinging. Ten swings should take approximately 16 seconds.

concept 1

Physical pendulum
Mass is distributed
Motion approximates SHM

equation 1

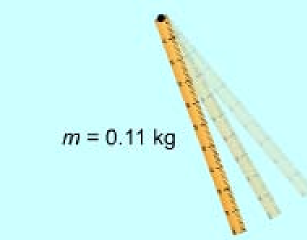


Physical pendulum period

$$T = 2\pi\sqrt{\frac{I}{mgh}}$$

T = period
I = moment of inertia
m = mass, g = acceleration of gravity
h = distance from pivot to center of mass

example 1



What is the period of the swinging meter stick?

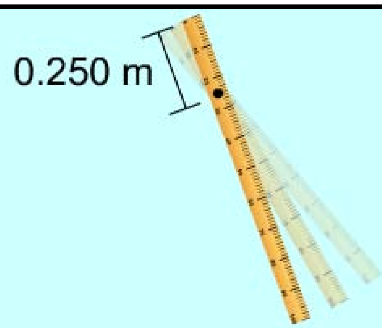
$$T = 2\pi\sqrt{\frac{I}{mgh}}$$

I = 1/3 × mass × length^2

$$T = 2\pi\sqrt{\frac{\frac{1}{3}(0.11\ \text{kg})(1.0\ \text{m})^2}{(0.11\ \text{kg})(9.80\ \text{m/s}^2)(0.50\ \text{m})}}$$

$$T = 2\pi\sqrt{0.068} = 1.6\ \text{s}$$

15.26 - Sample problem: meter-stick pendulum

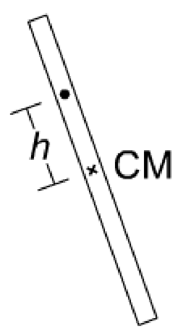


A meter stick swings as a pendulum, with an axis of rotation 0.250 meters from the end. What is its period?

Since the meter stick is a physical pendulum, its period depends on its moment of inertia. The meter stick is assumed to be uniform. In order to calculate its moment of inertia, we will treat it as a thin rod and use the parallel axis theorem.

Notice that you are not given the mass of the meter stick. It turns out it is not needed to solve the problem.

Diagram



We show the center of mass, labeled CM, and the distance from the center of mass to the axis of rotation, labeled h. The distance h will be needed to calculate the moment of inertia.

Variables

length of stick	L = 1.00 m
distance from axis to CM	h = 0.250 m
mass	m
moment of inertia	I

What is the strategy?

1. Use the parallel axis theorem to calculate the moment of inertia of the meter stick for the axis of rotation. This expression will involve the mass.
2. Use the moment of inertia to calculate the period. The mass will cancel out.

Physics principles and equations

The period of the physical pendulum depends on its moment of inertia, mass and the distance from the pivot to the center of mass.

$$T = 2\pi\sqrt{\frac{I}{mgh}}$$

The parallel axis theorem calculates the moment of inertia for an axis of rotation parallel to one passing through the center of mass.

$$I = I_{CM} + mh^2$$

For a thin rod, the moment of inertia for an axis of rotation passing through the center of mass is given by

$$I_{CM} = \frac{1}{12}mL^2$$

Step-by-step solution

We first calculate the moment of inertia of the meter stick about the indicated axis, using the parallel axis theorem. We do not know the mass, so we leave that variable in for now.

Step	Reason
1. $I = I_{CM} + mh^2$	parallel axis theorem
2. $I = \frac{1}{12}mL^2 + mh^2$	moment of inertia of thin rod about CM
3. $I = \frac{1}{12}(m\text{ kg})(1.00\text{ m})^2 + (m\text{ kg})(0.250\text{ m})^2$	enter values
4. $I = (0.146)m\text{ kg}\cdot\text{m}^2$	evaluate

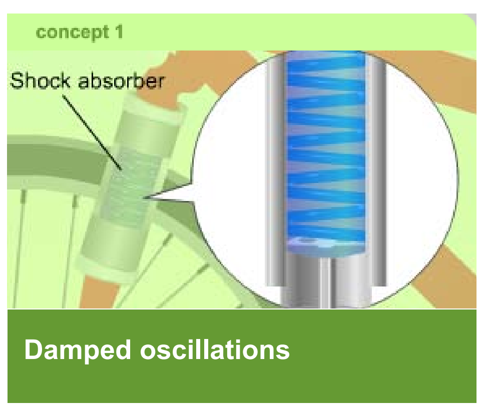
Now we use the moment of inertia to find the period.

Step	Reason
5. $T = 2\pi\sqrt{\frac{I}{mgh}}$	physical pendulum period
6. $T = 2\pi\sqrt{\frac{(0.146)m\text{ kg}\cdot\text{m}^2}{(m\text{ kg})(9.80\text{ m/s}^2)(0.250\text{ m})}}$	enter values
7. $T = 2\pi\sqrt{\frac{0.146\text{ m}^2}{(9.80\text{ m/s}^2)(0.250\text{ m})}}$	cancel m
8. $T = 1.53\text{ s}$	evaluate

In a previous section, we found that the period for a meter stick rotating at its end was a slightly longer 1.6 seconds.

15.27 - Damped oscillations

We have considered many types of oscillations, and up until now assumed the periodic motion continued without change. But most real-world oscillations are *damped*, which means they are subject to forces like friction that cause the amplitude of the motion to decrease over time.



Mountain bike shock absorbers provide an excellent demonstration of damped oscillations. A shock absorber often combines a spring with a sealed container of fluid. Shock absorbers lessen the jolts of a bumpy trail.

To explain this in more detail, let's consider what happens when a bike equipped with such a shock absorber hits a bump. The force from the bump compresses the spring, with the result that less of the force from the bump passes to the rest of the bike (and the rider).

The spring then supplies a restoring force. In the absence of any other force, the rider and bike would in principle then move forever in simple harmonic motion. However, inside a shock absorber, the spring moves a piston in a sealed cylinder of fluid. The fluid supplies what is called a *damping force*.

In Concepts 1 and 2, you see a diagram of this system. The fluid (typically oil) provides a force that opposes the motion of the piston. The damping force always opposes (resists) the motion of an object, which means sometimes it acts in the same direction as the restoring force (when the object moves away from equilibrium), and sometimes in the opposite direction (when the object moves toward equilibrium). At all times, however, it is opposing the motion.

Instead of moving in SHM, the system moves back to its equilibrium point and stops, or it may oscillate a few times with smaller and smaller amplitude before resting at its equilibrium point. The fluid "dampens" the motion, reducing the amplitude of the oscillations. The result is a relatively fast yet smooth return to the equilibrium position.

The resistive force of the fluid in a system like this is often proportional to the velocity, and opposite in direction. In Equation 1, you see the equation for the damping force. It equals the negative of b (the *damping coefficient*) times the velocity. (You may note that this is similar to the formula for air resistance, where the drag force depends on the square of the velocity.) The negative sign indicates that the damping force opposes the motion that causes it.

In Equation 2, you also see the equation for the net force F_N. The net force is the sum of the restoring forces and the damping force. (If you look at the equation, it may seem that two negatives combine to make a larger number, but the sign of the velocity is the opposite of the displacement as the system moves toward equilibrium.)

The graph in Equation 3 illustrates three types of damping. The blue line represents a *critically damped* system. The damping force is such that the system returns to equilibrium as quickly as possible and stops at that point.

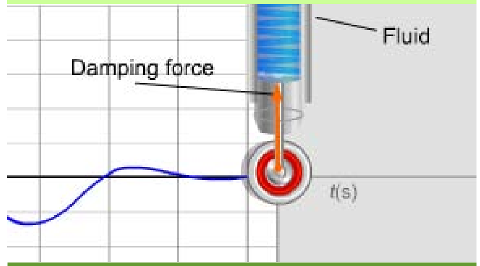
The green line represents a system that is *overdamped*. The damping force is greater than the minimum needed to prevent oscillations. The system returns to equilibrium without oscillating, but it takes longer to do so than a critically damped system.

The red line is a system that is *underdamped*. It oscillates about the equilibrium point, with ever diminishing amplitude.

With certain shock absorbers, the system can be adjusted, which means that the damping coefficient can be tuned based on rider preferences. Beginners often prefer an underdamped system. The bike bounces a bit but there is less of a "jolt" because the shock absorber acts more slowly. Advanced riders sometimes prefer a critically damped or overdamped "harder" ride, trading off a less smooth ride in exchange for regaining control of the bicycle more quickly.

Damping causes oscillations to diminish

concept 2



Damping force
Opposes motion
Often proportional to velocity

equation 1

Damping force

$$F_d = -bv$$

F_d = damping force
b = damping coefficient
v = velocity

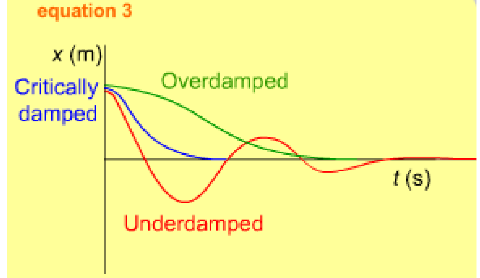
equation 2

Net force

$$\Sigma F = -kx - bv$$

ΣF = net force
k = spring constant
x = displacement
b = damping coefficient
v = velocity

equation 3



Types of damped harmonic motion
Critically damped: blue line
Overdamped: green line
Underdamped: red line

15.28 - Forced oscillations and resonance

Forced oscillation: A periodic external force acts on an object, increasing the amplitude of its motion.

External forces can dampen, or reduce, the amplitude of harmonic motion. For instance, in a mass-spring system, friction reduces the amplitude of the mass's motion over time. External forces can also maintain or increase the amplitude of an oscillation, counteracting damping forces.

Consider a child on a swing. Friction and air resistance are damping forces that reduce the amplitude of the motion. On the other hand, an external force like a person pushing, as you see in Concept 1, can increase the amplitude. When an external force increases the amplitude, *forced oscillation* occurs.

An external force that acts to increase the amplitude of oscillations is called a *driving force*. The driving force oscillates at a frequency called the *driving frequency*.

The natural frequency of a system is the frequency at which it will oscillate in the absence of any external force. Systems have natural frequencies based on their structure. The closer the driving frequency is to the natural frequency, the more efficiently the driving force transfers energy to the system, and the greater the resulting amplitude. This is why you push a child on a swing "in sync" with the swing's motion. The resulting phenomenon is called *resonance.* When the driving and natural frequencies are the same, the result is called *perfect resonance*.

There are several famous/infamous cases of forced oscillations and resonance. In Equation 1, you see a movie of the Tacoma Narrows Bridge. A few months after it was built in 1940, strong winds caused the bridge to oscillate at its natural frequency, and the amplitude of the oscillations increased over time until the bridge collapsed. The precise cause of the collapse is a matter of some debate, but the resonant oscillations played a large part.

The Bay of Fundy in Nova Scotia provides another famous example. The tides vary greatly in the bay with the water level changing by as much as 16 meters. One reason for the dramatic tides is that the natural frequency of the bay, the time it takes for a wave to go from one end to the other, is close to the driving frequency of the tide cycle, which is about 12.5 hours.

As a third example, the natural frequency of one- to three-story buildings is close to the driving frequency supplied by some earthquakes, which is why these buildings (very common in San Francisco) often sustain the heaviest damage during quakes.

In Equation 2, you see a graph called a *resonance curve*. It is a graph of amplitude versus frequency for a system that has both a damping force and an external driving force. We call the natural (angular) frequency ω_n and use ω to indicate the driving frequency. As the driving frequency ω approaches the natural frequency ω_n, the amplitude increases dramatically.

Natural frequencies can be "natural," but in some cases they can also be controlled. Electric circuits, such as those used to tune radios to stations of different frequencies, are designed so that humans can change the natural frequency of the circuit. As you turn the radio dial, you are changing the natural frequency of the circuit. It then "tunes in" a driving frequency from a radio station that matches the natural frequency of the circuit. These concepts have entered everyday language. People say that "an idea resonates with me." Such everyday speech is good physics; they mean the "driving frequency" of the idea is close to the "natural frequency" of their own beliefs.

concept 1

Forced oscillations
External force in direction of motion
Amplitude increases

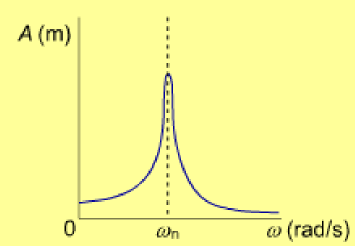
equation 1

Resonance

$$\omega \approx \omega_n$$

ω = driving frequency
ω_n = natural frequency

equation 2



Frequency and amplitude
Amplitude increases as ω approaches ω_n

15.29 - Gotchas

To calculate the amplitude of an object moving in SHM, measure the difference between two successive peaks of its graph. No, that is the period you just measured. The amplitude is the height of a peak of the graph above the horizontal (time) axis.

The slope at any point on the displacement graph of an object in SHM is its velocity. Yes, you are correct. This is a point that is true of any displacement graph, not just an SHM graph.

15.30 - Summary

Simple harmonic motion (SHM) is a kind of repeated, consistent back and forth motion, like the swinging of a pendulum. It is caused by a restoring force that varies linearly with displacement.

The displacement associated with such motion can be described with a sinusoidal function, typically a cosine. The displacement is zero at equilibrium and maximum at the extreme positions.

Just as with other types of repetitive motion, the period of SHM is the amount of time required to complete one cycle of motion. The frequency is the number of cycles completed per second. It is the reciprocal of the period. The unit of frequency is the hertz (Hz), equal to one inverse second.

Angular frequency is the frequency measured in radians per second. It is represented by the Greek letter ω and is seen in the function for harmonic motion.

If the object in simple harmonic motion is a mass on a spring, the spring constant and the mass determine the angular frequency.

The amplitude of harmonic motion is the maximum displacement from equilibrium. It is represented by A and appears as the coefficient of the cosine in the displacement function for SHM.

The phase constant, φ, specifies the position at the zero time, effectively shifting the graph to the left or right.

The velocity and acceleration functions for SHM are also sinusoidal. The maximum velocity occurs at equilibrium, and it is zero at the extremes. Acceleration is the opposite: zero at equilibrium and maximum at the extremes. These relationships follow from the general nature of velocity as the instantaneous slope of the displacement graph, and acceleration as the slope of velocity.

As an object like a mass on a spring moves in SHM, its total energy remains constant, although it transforms from potential to kinetic energy and back continuously. If you stretch the spring and release it to begin the motion, the amount of work you do on the spring is the amount of potential energy you have stored in it. All the energy is potential energy at the extremes, and kinetic energy at the equilibrium position.

A simple pendulum displays simple harmonic motion in its angular displacement, provided that the amplitude of the motion is small. Instead of a restoring force, there is a restoring torque due to gravity. The period of a pendulum depends upon the length of the pendulum and the acceleration of gravity.

The simple pendulum is a special case of the more complicated physical pendulum. In general, the period of a physical pendulum depends upon its moment of inertia, mass, and the distance from the pivot point to its center of mass, as well as the acceleration of gravity.

Sometimes a damping force opposes oscillatory motion. A typical damping force is proportional to the velocity of the object, which changes with time.

A force that acts **with** the restoring force can maintain or increase the amplitude of oscillations. Forced oscillations occur when such a driving force is present. The natural frequency of a system is the frequency at which it will oscillate in the absence of external force. As the frequency of the driving force approaches the natural frequency, energy is transferred more efficiently and the system's oscillation amplitude increases. When these frequencies are approximately equal, resonance occurs.

Equations

$$x(t) = A\cos(\omega t + \varphi)$$

$$f = 1/T$$

$$\omega = 2\pi f$$

$$v(t) = -A\omega\sin(\omega t + \varphi)$$

$$a(t) = -A\omega^2\cos(\omega t + \varphi)$$

$$\omega = \sqrt{\frac{k}{m}}$$

$$T = 2\pi\sqrt{\frac{m}{k}}$$

$$PE = ½\, kx^2$$

$$PE = ½\, kA^2\cos^2\omega t$$

$$KE = ½\, mA^2\omega^2\sin^2\omega t$$

$$TE = ½\, kA^2$$

Chapter 15 Problems

Chapter Assumptions

The general form of the equation of motion for an object in SHM is $x(t) = A\cos(\omega t + \varphi)$.

Conceptual Problems

C.1 A bouncing ball returns to the same height each time. Is this an example of simple harmonic motion? Explain your answer.

◯ Yes ◯ No

C.2 Consider the displacement, velocity, and acceleration vectors of an object moving in SHM. Which pair of these vectors always point in opposite directions?

- [ ] Displacement
- [ ] Velocity
- [ ] Acceleration

C.3 In old pocket watches, a balance wheel acts as a torsional pendulum, rotating with a fixed period. If a pocket watch is running slow, the period of the balance wheel is too long. Would you add or remove mass from the outer edge of the balance wheel to correct it?

◯ Remove mass
◯ Add mass

C.4 For a physical pendulum, what happens to the period as the pivot point gets very close to the center of mass? Justify your answer using the equation for the period of a physical pendulum.

i. The period gets longer
ii. The period shortens
iii. The period says the same

C.5 What are the units for the damping coefficient constant?

◯ N/m
◯ kg/s
◯ kg/s^2

Section Problems

Section 0 - Introduction

0.1 Using the simulation in the interactive problem in this section, answer the following questions. (a) If you increase the amplitude, does the period increase, decrease, or stay the same? (b) What does the shape of the curve look like?

(a) i. Increase
ii. Decrease
iii. Stay the same

(b) i. Line
ii. Parabola
iii. Sinusoidal function
iv. Circle

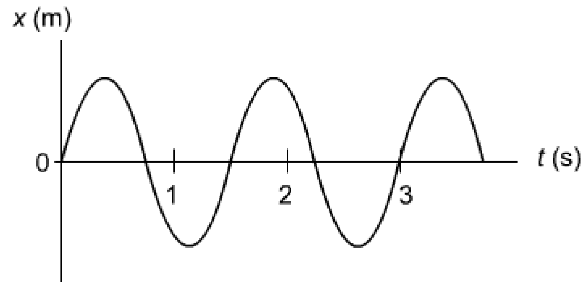
Section 3 - Period and frequency

3.1 Consider the minute hand on a clock. (a) Compute the frequency of its motion in cycles per second. State your answer to three significant digits. (b) Do the same for the hour hand.

(a) __________ Hz
(b) __________ Hz

3.2 A graph of the displacement of an object moving in SHM is shown. Determine the frequency of the object's motion. (Assume you can read the graph points to two significant figures.)

__________ Hz



Section 4 - Angular frequency

4.1 What is the angular frequency of the second hand on a clock? (State your answer using three significant figures.)

__________ rad/s

4.2 A potter's wheel rotates with an angular frequency of 1.54 rad/s. What is its period?

__________ s

Section 5 - Amplitude

5.1 What is the amplitude of an object moving in SHM if its displacement in meters is described by:

(a) $x(t) = 5\cos(t - \pi/2)$

(b) $x(t) = 4\cos(2\pi t) - \cos(2\pi t)$

(c) $x(t) = 4\cos^2(\pi t) - 4\sin^2(\pi t)$

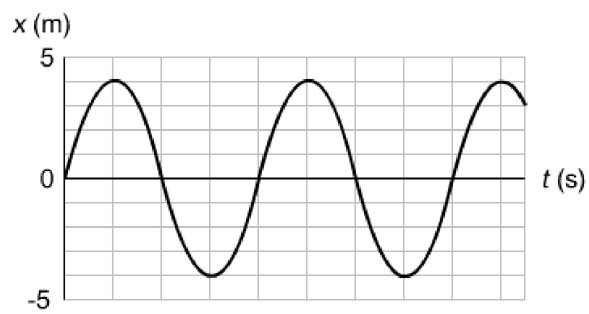
(a) __________ m

(b) __________ m

(c) __________ m

5.2 The displacement of an object moving in SHM is graphed as shown. What is its amplitude of motion? (Assume you can read the graph points to two significant figures.)

__________ m



5.3 An object moving in SHM has an amplitude of 3.5 m and a period of 4.0 s. Which of the following equations could describe its displacement over time?

- ○ $x(t) = 2.0\cos(3.5t)$
- ○ $x(t) = 3.5\cos(t + 2.0)$
- ○ $x(t) = 3.5\cos((\pi/2)t + 2.0)$
- ○ $x(t) = 3.5\cos(2.0\pi t)$

5.4 The displacement of an object in meters is described by the function $x(t) = 4.4\cos(7.4\pi t + \pi/2)$ where t is measured in seconds. What are the (a) amplitude, (b) frequency, (c) period of the object's motion?

(a) __________ m

(b) __________ Hz

(c) __________ s

Section 6 - Interactive problem: match the curve

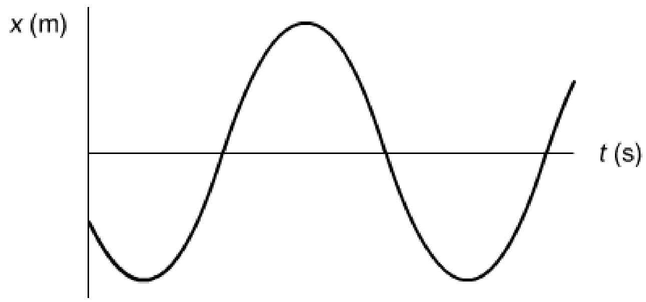
6.1 Using the simulation in the interactive problem in this section, what (a) amplitude and (b) period should be used to match the graph?

(a) __________ m

(b) __________ s

Section 7 - Phase and phase constant

7.1 The displacement graph of an object moving in SHM is shown. Which of the following describe the phase φ for the object's motion?



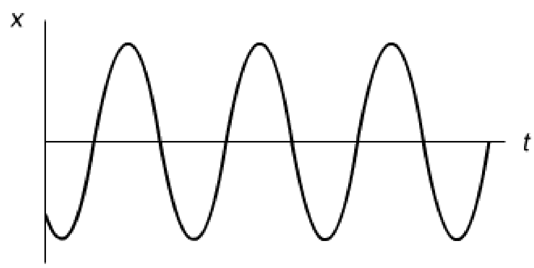
- ○ $0 < \varphi < \pi/2$
- ○ $\pi/2 < \varphi < \pi$
- ○ $\pi < \varphi < 3\pi/2$
- ○ $3\pi/2 < \varphi < 2\pi$

7.2 An object moves in SHM with amplitude 4.37 m. The phase constant of the function describing its motion is 1.33 rad. What is the object's displacement at time t = 0 s?

__________ m

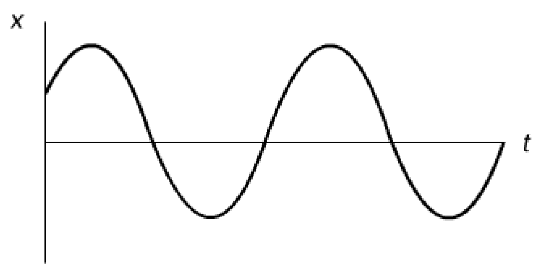
7.3 The graph of an object moving in SHM is shown. The amplitude of its motion is 2.7 m and at time t = 0 s, its displacement is −1.6 m. What is the phase constant φ?

__________ rad



7.4 The graph of an object moving in SHM is shown. The amplitude of its motion is 3.6 m and at time t = 0 s, its displacement is 2.6 m. What is the phase constant φ? Give a positive answer between 0 and 2π.

__________ rad



Section 8 - Sample problem: graph equation

8.1 An object moving in SHM has its most negative displacement at time t = 0 s. If the amplitude of the motion is 7.8 m and the angular frequency is 5.6 rad/s, which of these equations describes the motion?

- ○ $x(t) = 7.8 \cos (5.6t)$
- ○ $x(t) = 7.8 \cos (5.6\pi t)$
- ○ $x(t) = 7.8 \cos (5.6t + \pi/2)$
- ○ $x(t) = 7.8 \cos (5.6t + \pi)$

8.2 An object moving in SHM has zero displacement at time t = 0 s, and positive displacement a moment after. If the amplitude of the motion is 4.8 m and the frequency is 4.6 Hz, which of these equations describes the motion?

- ○ $x(t) = 4.8 \cos (9.2\pi t + 3\pi/2)$
- ○ $x(t) = 4.8 \cos (9.2\pi t + \pi/2)$
- ○ $x(t) = 4.8 \cos (9.2t + 3\pi/2)$
- ○ $x(t) = 4.8 \cos (4.6t + \pi/2)$

8.3 An object moving in SHM has maximum (positive) displacement at time t = 0 s. If the amplitude of the motion is 2.1 m and the period is 6.4 s, what is the displacement at time t = 2.8 s?

__________ m

8.4 An object moving in SHM has zero displacement at time t = 0 s, and negative displacement a moment after. If the amplitude of the motion is 6.8 m and the frequency is 4.6 Hz, what is the displacement at time t = 3.4 s?

__________ m

Section 9 - Velocity

9.1 In a car engine, a piston moves in SHM with an amplitude of 0.410 m. The engine is running at 2400 rpm, which is an angular frequency of 251 rad/s. What is the maximum speed of the piston?

__________ m/s

9.2 The equation for the displacement in meters of an object moving in SHM is $x(t) = 1.50 \cos(4.20t)$ where t is in seconds. (a) What is the maximum speed of the object? (b) At what time does it **first** reach the maximum **speed**?

(a) __________ m/s
(b) __________ s

Section 11 - Acceleration

11.1 A ball on a spring moves in SHM. At time t = 0 s, its displacement is 0.50 m and its acceleration is −0.72 m/s^2. The phase constant for its motion is 0.84 rad. What is the ball's displacement at t = 3.4 s?

__________ m

11.2 A platform moves up and down in SHM, with amplitude 0.050 m. Resting on top of the platform is a block of wood. What is the shortest period of motion for the platform so that the block will remain in constant contact with it?

__________ s

Section 13 - Summary of simple harmonic motion

13.1 A particle moves in SHM, with displacement defined by the equation $x(t) = 0.0018 \cos(2\pi t)$, where x is measured in meters and t in seconds. (a) What is the particle's maximum speed? (b) What is its maximum acceleration? (c) What is its acceleration when $x = 0.0013$ m?

(a) __________ m/s
(b) __________ m/s^2
(c) __________ m/s^2

13.2 A block attached to a spring moves in SHM on a frictionless surface. The acceleration of the block is given by the equation $a(t) = -3.6\, x(t)$, where x is measured in meters and t in seconds. (a) What is the angular frequency of the block's motion? (b) When the block has maximum acceleration, its displacement is −2.3 m. What is the amplitude of the block's motion?

(a) __________ rad/s
(b) __________ m

13.3 A particle is moving in SHM. Its maximum acceleration is 2.27 m/s^2 and its maximum velocity is 1.39 m/s. What is the period of its motion?

__________ s

13.4 The acceleration of an object moving in SHM is defined by the equation $a(t) = -3.72x(t)$, where x is measured in meters and t in seconds. At time $t = 0$ s, the object has its maximum positive displacement. (a) What is the first time after $t = 0$ s that the object is at the equilibrium position (zero displacement)? (b) In what direction is the object moving at that time?

(a) __________ s
(b) ○ Positive direction
○ Negative direction

Section 14 - Simple harmonic motion and uniform circular motion

14.1 A particle moves in a circle of radius 0.45 m at a constant speed, completing one revolution every 1.2 s. The particle is on the positive x axis at time t = 0 s. Write an equation for the y displacement of the particle as a function of time.

○ $y(t) = 0.45 \cos((2\pi/1.2)t)$
○ $y(t) = 0.45 \sin((2\pi/1.2)t)$
○ $y(t) = 0.45 \cos(1.2t)$
○ $y(t) = 0.45 \sin(1.2t)$

14.2 The text states an equation for the x displacement of an object in uniform circular motion: $x(t) = A \cos \omega t$. Write a corresponding function for the y displacement, also using the cosine function.

○ $y(t) = 2A \cos \omega t$
○ $y(t) = A \cos(\omega t - \pi)$
○ $y(t) = A \cos(\omega t - \pi/2)$
○ $y(t) = A \cos(\omega t - \pi/4)$

Section 15 - Period, spring constant, and mass

15.1 A block with mass 0.55 kg on a frictionless surface is attached to a spring with spring constant 41 N/m. The block is pulled from the equilibrium position and released. What is the period of the system?

__________ s

15.2 A 0.683 kg mass moves in SHM at the end of a spring. It takes 1.41 s to move from the position with the spring fully extended to the position with the spring fully compressed. What is the spring constant?

__________ N/m

15.3 A block with mass 0.382 kg is attached to a horizontal spring with spring constant k = 1.28 N/m on a frictionless surface. The block is pulled 0.753 m from equilibrium and released. (a) What is the amplitude of the block's motion? (b) What is its period? (c) How long after release does the block take to first return to its equilibrium position? (d) What is its speed at that position?

(a) __________ m
(b) __________ s
(c) __________ s
(d) __________ m/s

15.4 A 1.20 kg mass is attached to the end of a spring of unknown spring constant. The spring is compressed a distance of 0.300 meters, and when it is released, the mass oscillates horizontally on a frictionless surface. When the mass is 0.100 m from equilibrium, it is seen to be moving at a speed of 2.00 m/s. (a) What is the amplitude of the motion? (b) Find the spring constant, k. (c) What is the maximum speed of the mass?

(a) __________ m
(b) __________ N/m
(c) __________ m/s

15.5 A car has a mass of 1550 kg. Its four passengers have a combined mass of 322 kg. The car's suspension has four identical springs. The suspension frequency for a comfortable ride is 1.300000111474276 Hz. Assume the load is distributed evenly. What should the spring constant for each spring be?

__________ N/m

15.6 A spring of unknown spring constant is attached to a 2.30 kg mass. The mass is pulled horizontally outward by a distance "A" from equilibrium, then released at t = 0. At t = 0.200 seconds, the mass first reaches the equilibrium position, where it is seen to be moving at a velocity of v = −0.950 m/s. (a) What is the period of the motion? (b) What is the frequency of the motion? (c) Calculate the spring constant of the spring. (d) Calculate the amplitude of motion, A.

(a) __________ s
(b) __________ Hz
(c) __________ N/m
(d) __________ m

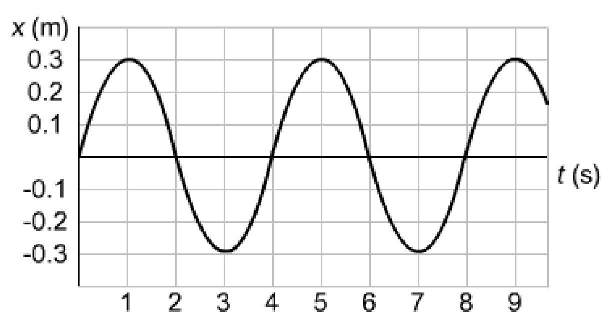
15.7 A large cube made of cork floats in a swimming pool, with one face parallel to the water surface. The cube is uniform in density, with mass 11.6 kg and each edge 0.381 m in length. The top surface of the cube is briefly tapped, causing it to oscillate in the vertical direction in SHM. The cube is damp, of course, but its motion is not damped. What is the period of the cube's motion? Use a value of 997 kg/m^3 for the density of the water, an appropriate value for the water in a swimming pool at 25°C.

__________ s

Section 16 - Interactive problem: match the curve again

16.1 A block at the end of a horizontal spring moves in SHM as shown by the graph. The mass of the block is 0.38 kg. Assume you can read the graph points to two significant figures. (a) What is the spring constant? (b) What is the block's maximum acceleration?

(a) __________ N/m
(b) __________ m/s^2



16.2 A block on a spring moves in simple harmonic motion. The mass of the block is 0.30 kg and the spring constant is 0.80 N/m. If the block is at its maximum positive displacement of 0.70 m at time t = 0 s, draw a graph of the block's displacement over the first three seconds.

16.3 Using the simulation in the interactive problem in this section, what should the (a) amplitude and (b) the mass of the puck be in order to match the graph?

(a) __________ m
(b) __________ kg

Section 18 - Work and the potential energy of a spring

18.1 A motor pulls a spring 0.73 m away from its equilibrium position, doing 3.5 J of work on the spring in the process. What is the spring constant?

__________ N/m

18.2 A spring with a spring constant of 45 N/m is pulled 1.4 m away from its equilibrium position. How much potential energy is stored in the spring?

__________ J

Section 19 - Total energy

19.1 A block with mass 0.67 kg, resting on a horizontal frictionless surface, is attached to the end of a spring with spring constant 49 N/m. The block is released from rest at a distance 0.035 m from the equilibrium position. What is the total energy of the system?

__________ J

19.2 A block of mass 0.35 kg oscillates in SHM on a spring with an amplitude of 0.96 m. The maximum acceleration of the block is 2.7 m/s^2. What is the total energy of the system?

__________ J

19.3 An object is moving in simple harmonic motion with amplitude 0.28 m and total energy 2.5 J. By some means its amplitude is increased to 0.34 m. What is the total energy of the system now?

__________ J

19.4 A block and spring system oscillating in SHM has a maximum speed of 1.18 m/s and a total energy of 1.71 J. What is the mass of the block?

__________ kg

Section 20 - Interactive checkpoint: spring energy and period

20.1 A 4.6 kg block is attached to a spring with a spring constant of k = 6.3 N/m. It oscillates in SHM with an amplitude of 1.8 m. What is the maximum speed of the block?

__________ m/s

20.2 A block on a spring moves in SHM with amplitude 0.23 m and period 3.8 s. The mass of the block is 2.1 kg. What is the total energy of the system?

__________ J

Section 21 - Sample problem: falling block on a spring

21.1 A 2.7 kg block hangs from a vertical spring whose upper end is fixed. The spring constant is 93 N/m. Define the *PE* of the system to be zero when the block is at the equilibrium position. The block is then set into motion and oscillates with an amplitude of 3.8 m. When the block is at its lowest position, what is the potential energy stored in the spring?

__________ J

21.2 A 3.5 kg block oscillates in SHM from the end of a vertical spring with spring constant 88 N/m. The block's maximum speed is 12 m/s. What is the block's amplitude of motion?

__________ m

21.3 An object is hung at the end of a vertical unstretched spring, and released. It falls 0.347 m before reversing direction. What is its period of motion?

__________ s

21.4 A 2.7 kg block hangs at the end of a spring. When the block is at rest, the spring is stretched 0.38 m away from its equilibrium position. The block is set into simple harmonic motion and oscillates with an amplitude of 0.24 m. What is the total energy of the block/spring system?

__________ J

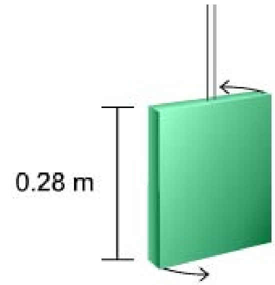
Section 22 - A torsional pendulum

22.1 An irregularly-shaped 1.4 kg object is suspended from a wire with a known torsion constant of 0.49 N·m/rad. The object's period is 1.2 seconds. What is the object's moment of inertia for rotations about this axis?

__________ kg · m2

22.2 A thin square slab of material is suspended "on edge" at the end of a torsion pendulum, so that the axis of rotation passes through the center of the square, parallel to an edge. The mass of the slab is 0.78 kg and the length of an edge is 0.28 m. The torsion constant of the wire is 6.2 N·m/rad. What is the period of motion when the slab oscillates?

__________ s



Section 23 - A simple pendulum

23.1 You need to know the height of a room, but you have no tape measure. You fasten one end of a string to the ceiling of the room, and tie a small rock at the other end so it almost touches the floor. You start this simple pendulum swinging slightly, and measure its period, which is 3.56 seconds. How tall is the room?

__________ m

23.2 On the moon of a distant planet, an astronaut measures the period of a simple pendulum, 0.85 m long, and finds it is 4.7 seconds. Back on Earth, she could throw a rock 13 m straight up (while wearing her spacesuit). With the same effort, how far up can she throw the same rock at her present location? Ignore the effects of air resistance.

__________ m

23.3 It is the year 2305 and the tallest structure in the world has an insane height of 3.19×10^6 m above the surface of the Earth. A pendulum clock that keeps perfect time on the surface of the Earth is placed at the top of the tower. How long does the clock take to register one elapsed hour? The radius of the Earth is 6.38×10^6 m and its mass is 5.97×10^{24} kg.

__________ minutes

Section 24 - Interactive problem: a pendulum

24.1 Using the simulation in the interactive problem in this section, what is the length of string needed to achieve the desired period for the pendulum?

__________ m

Section 25 - Period of a physical pendulum

25.1 A flat circular disk with diameter 0.36 m and mass 0.43 kg is suspended so it can swing freely from a pivot 0.12 m from its center. (The axis of rotation is perpendicular to the plane of the disk.) What is its period of oscillation?

__________ s

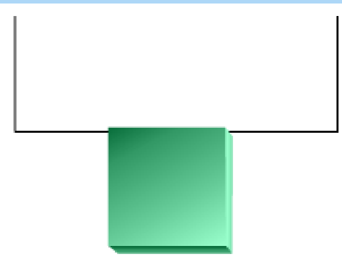
25.2 A small hole is drilled in a meter stick is to act as a pivot. The meter stick swings in a short arc as a physical pendulum. How far from the center of mass should the pivot point be for a period of 2.00 seconds?

__________ m

Section 26 - Sample problem: meter-stick pendulum

26.1 A thin square slab of material is suspended so that it can rotate freely around one of its edges, which is parallel to the ground. The slab is uniform in density and has mass 0.78 kg. The length of each side is 0.67 m. What is the slab's period?

__________ s



26.2 A disk with diameter 4.30 m is allowed to oscillate freely as a physical pendulum about an axis that is perpendicular to its surface, and located some distance from its center. If its period of oscillation is 3.50 s, what is the minimum possible distance of the axis from the center of the disk?

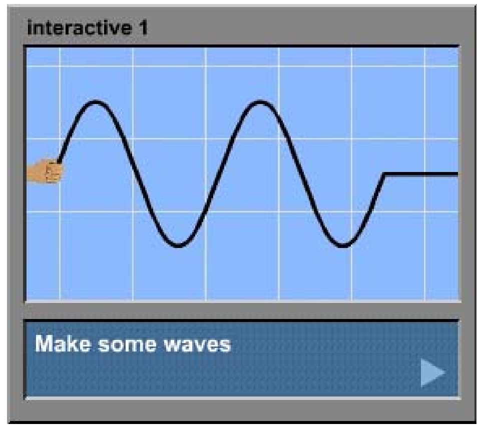
__________ m

26.3 A solid spherical ball of mass 0.227 kg is suspended from a thin rod of negligible mass to form a pendulum. The diameter of the ball is 0.214 m and the length of the rod is 0.125 m. (a) If you considered this to be a simple pendulum, with length equal to the rod length plus the radius of the ball, what would its period be? (b) What is the actual period of the pendulum?

(a) __________ s
(b) __________ s

chapter 16 Wave Motion

16.0 - Introduction



Waves can be as plain to see as the ripples in a pond or as invisible as the electromagnetic waves emanating from a cellular phone. Mechanical waves, like those in a pond, require a medium in order to propagate. Electromagnetic waves − including radio waves and light − require no medium and can travel in the near vacuum of space. Electromagnetic waves rely on the interaction of electric and magnetic fields to propagate through space.

In this chapter, we focus on mechanical waves. These are waves in which a vibration causes a disturbance to travel through a medium. You are familiar with a variety of mechanical waves: water waves in the ocean, sound waves in the air, or waves along a string if you shake an end up and down. These waves exist due to the movement of particles that make up a medium, such as water molecules in the ocean or gas molecules in the air.

Waves carry energy from place to place: a relatively small amount with a sound wave, a much larger amount with a tsunami wave.

Although many mechanical waves travel, sometimes across great distances, there is no net movement of the medium through which they propagate. The 15th century Italian scientist and artist Leonardo da Vinci described this key attribute when he said: “It often happens that the wave flees the place of its creation, while the water does not.”

Use the simulation to the right to begin your exploration of waves. It consists of a string stretched across the screen. A hand on the left holds the string. By shaking the hand up and down, you can generate a variety of waves in the string.

When you open the simulation, press GO to send a wave down the string. You will see the hand begin to shake the string, causing a wave to travel from left to right.

The control panel has two input gauges that allow you to vary the amplitude and frequency of the wave. As you may remember from your study of simple harmonic motion, amplitude is the maximum displacement of a wave from equilibrium. Frequency is the number of cycles per second. You can vary these parameters and observe changes in the shape of the wave. Also in the control panel is an output gauge that displays the wavelength, the distance between successive peaks of the wave.

The string’s tension and other properties remain constant.

When you run the simulation, make sure you observe the differences between a wave with higher frequency and one with lower frequency. This is an important fundamental characteristic of a wave.

Then try three quick experiments. First, does changing the frequency of a wave also change its wavelength? Change the frequency and observe what happens to the wavelength.

Second, does changing the frequency result in any change in amplitude? Again, you can vary the frequency and note any change.

Finally, as you change the frequency and amplitude, does the wave travel down the string any faster or slower? For example, does a wave with a very large amplitude travel noticeably faster than one with a very small amplitude?

The simulation is intended to let you conduct a preliminary exploration of topics that will be presented in this chapter. Answer what questions you can above and then proceed to the rest of the chapter, which covers the topics in more depth.

16.1 - Mechanical waves

concept 1

Mechanical waves
Disturbances in a medium

Mechanical waves: Vibrations in a medium.

Oceans, the rolling wave of a crowd in a sports stadium, the back and forth vibrations in a Slinky®: These are a few of the many kinds of waves you can see. Some mechanical waves are invisible to the eye but detectable by the ear, such as the sound waves generated by musical instruments.

Mechanical waves are vibrations in a medium, traveling from place to place without causing any net movement of the medium. You may be familiar with “the wave” in a football or baseball stadium. The wave travels around the stadium, the result of spectators standing and then sitting in a rolling succession. As the fans oscillate up and down, they create what is called a *disturbance* or *waveform*. The location of the disturbance changes as the wave moves through the stadium, but the wave’s medium, the crowd, stays put.

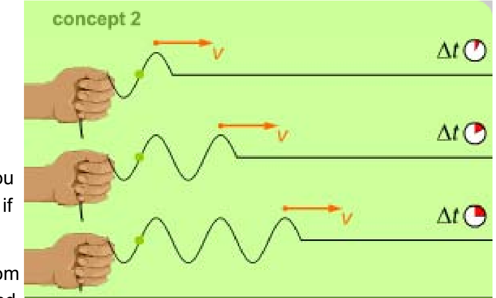
A wave in a stadium is a useful example, but it is not a true mechanical wave. Mechanical waves, such as a wave in a string, result from an initial force (a vibration up and down or to and fro) followed by a continuing sequence of interactions between particles in the medium. In a stadium wave, the particles of the medium (the people) do not typically exert physical forces on one another to propagate the wave (since peer pressure is not a physical force).

Mechanical waves have some common properties. First, they require a physical medium, such as air, a string or a body of water. Mechanical waves cannot move through a vacuum.

Second, mechanical waves require a driving excitation to get the wave started. The vibrations then propagate, via interactions between particles, through the medium.

In this fashion, waves transfer energy from place to place. When you hear a sound, you are hearing energy that has been transferred by a wave through air or water (or even, if you are listening for buffalo, the ground).

All of the waves in this chapter are *traveling waves* in which the disturbance moves from one point to another. Concept 2 shows a wave moving down a string, caused by a hand shaking the string. The illustration shows three successive moments in time. You can track the position of the first peak as it moves down the string over time. It moves with a constant speed v. The other peaks also move down the string with the same speed.



Traveling waves
Vibrations that travel through a medium

The peaks do not move through the medium in all waves. In what are called standing waves, the locations in the medium where peaks and troughs (the "low" parts of the wave) occur are fixed. These waves, caused by the reflection or interaction of traveling waves, are discussed in a later chapter.

16.2 - Transverse and longitudinal waves

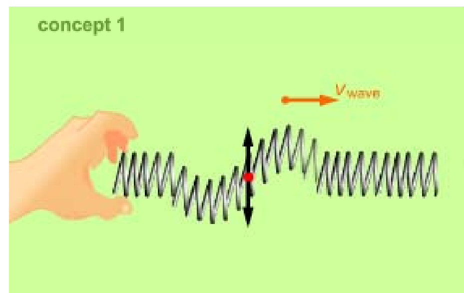
Transverse wave: Particles in a medium vibrating perpendicular to the direction the wave is traveling.

Longitudinal wave: Particles in a medium vibrating parallel to the direction the wave is traveling.

Waves can be classified by the relationship between their direction of travel, and the direction of the motion of the particles in the medium.

Imagine that two people stretch a Slinky between them, and one shakes the Slinky up and down. This causes a wave to move along the Slinky, as shown in Concept 1. The wave moves to the right with a velocity called v_{wave} in the diagram.

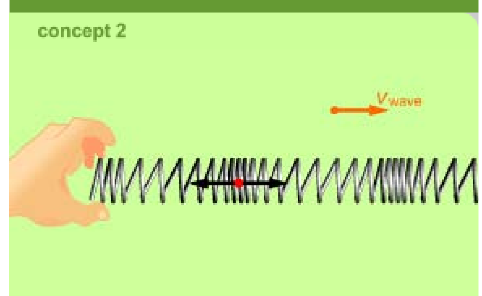
Although the wave moves to the right, the particles that make up the medium move up and down. An individual particle of the Slinky is highlighted in red in the diagram to the right, and its movement is shown with the vertically directed arrows. The direction of the wave is **perpendicular** to the motion of the particles of the medium. This type of wave is called a *transverse wave*.



Transverse waves
Particles vibrate perpendicular to direction of wave

Many types of mechanical waves are transverse waves, including those caused by shaking a Slinky up and down, the vibrations of a violin string, and certain types of earthquake waves.

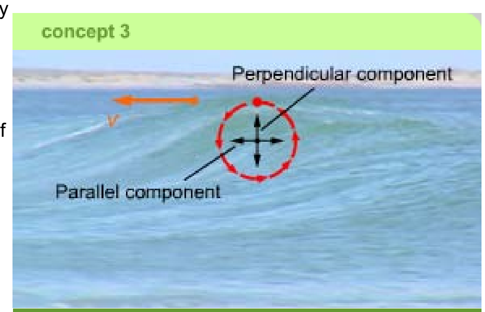
Now imagine that instead of shaking the Slinky up and down, a person pulls the Slinky to the left and then pushes it to the right, as shown in Concept 2. This causes the spring to be stretched and then compressed.



Longitudinal waves
Particles vibrate parallel to direction of wave

This disturbance again travels horizontally along the Slinky, and again we show its velocity as v_{wave}. The wave consists of regions in which the coils of the spring are tightly packed, followed by regions in which the coils are widely spaced. A particle of the Slinky, again marked with a red dot, oscillates horizontally, **parallel** to the direction the wave is traveling. This type of wave is called a *longitudinal wave*.

Sound is a longitudinal wave that consists of alternate compressions and rarefactions of air. Individual air particles oscillate back and forth, and a sound wave travels through the air, where it can be detected by a sophisticated instrument: the human ear.



Some waves both transverse and longitudinal
Water waves

In both transverse and longitudinal waves, the particles do move, but there is no **net** motion of the particles after each cycle. A particle moves up and down, or back and forth, but it returns to its initial position. It oscillates like a mass attached to a spring.

A single source of vibration, such as an earthquake, can create both transverse and longitudinal waves. In an earthquake, the longitudinal waves (*P waves*, for primary waves) travel at about 8 km/s, while the transverse waves (*S waves,* for secondary waves) are slower, moving at about 5 km/s. By noting when each type of wave arrives at a given seismographic station, a seismologist can determine the distance of the earthquake from that station. Using data from several stations, the seismologist can triangulate the location of the earthquake's epicenter.

· Wave travels horizontally
· Particle motion perpendicular and parallel

The motion of the particles that make up a wave can be complex, with ocean waves serving as one example. You can see this in Concept 3, where the wave moves to the left and the water molecules near the surface move in circles. This means the molecules' motion involves vertical and horizontal components. Their motion is both perpendicular and parallel to the direction of the wave. An ocean wave displays both longitudinal and transverse properties.

16.3 - Periodic waves

Wave pulse: A single disturbance caused by a one-time excitation.

Periodic wave: A continuing wave caused by a repeated vibration.

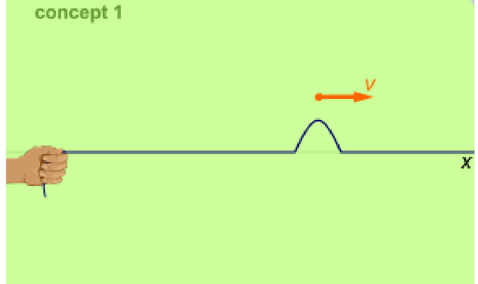
Two types of transverse waves are shown to the right. Concept 1 shows a single wave pulse in a string. The hand shakes up and down once and the wave pulse moves from left to right along the string.

Concept 2 shows a periodic wave. The hand moves continuously up and down. In a periodic wave, each particle in the string moves through a repeated cycle of rising to a peak, falling to a trough, and then returning again to a peak. The procession of wavefronts moving down the string is called a *wave train*.

If you observe a particular crest in the periodic wave, it will move horizontally along the string over time. This is more apparent in an animation. In a static diagram, the wave can appear to be stationary, though it is moving down the string as the velocity vector indicates. Click on the illustration to see an animation.

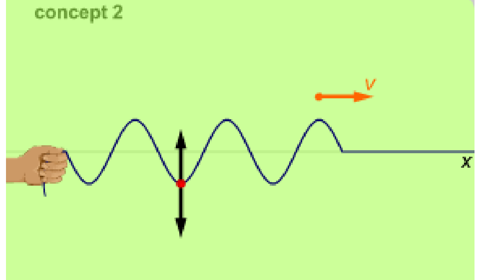
In this chapter, we focus on waves in which the particles are vibrating in simple harmonic motion. This vibration will be caused by something (in this case, a hand) moving or vibrating in simple harmonic motion. The result is the type of sinusoidal wave you see to the right.

concept 1



Wave pulse
Caused by a single up and down motion

concept 2



Periodic wave
Continuing wave caused by a repetitive vibration

16.4 - Amplitude

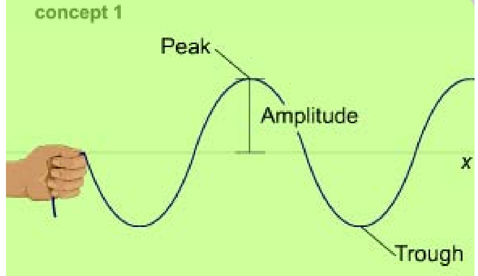
Amplitude: The maximum displacement of a particle in a wave from its equilibrium position.

Several terms discussed in earlier topics such as simple harmonic motion also apply to waves, including *amplitude*.

At the right, you see a transverse wave caused by a hand shaking a string. The amplitude of the wave is the distance between a particle at its maximum displacement − a peak or trough − and the particle at its rest or equilibrium position. The horizontal line in the diagram is the equilibrium position for the particles in the string. The amplitude by convention is positive. Since amplitude is a distance, it is measured in meters.

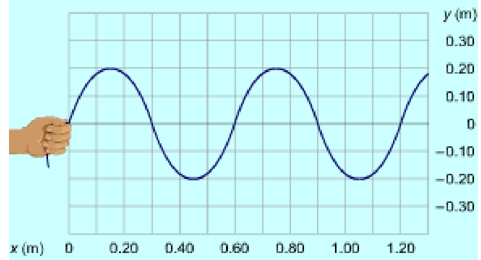
A wave's amplitude is related to the energy it carries. Waves with greater amplitude carry more energy. You can experience this relationship at the beach; you may barely notice a small-amplitude wave crashing into you, while a large-amplitude wave may knock you off your feet!

concept 1



Amplitude
Distance between rest point and maximum displacement
· Height of peak

example 1



What is the amplitude of this wave?

A = 0.20 m

16.5 - Wavelength

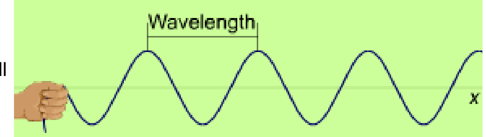
Wavelength: The distance between adjacent peaks.

The wavelength of a wave is the distance between adjacent peaks in the wave. This will be the same distance as that between adjacent troughs, or any two successive points on the wave with the same vertical displacement and direction of particle motion.

The Greek letter lamda (λ) represents wavelength. At the right, you see the wavelength measured for a transverse periodic wave in a string. The unit for wavelength is the meter.

Also on the right is a table that shows the wavelengths of a variety of waves.

concept 1



Wavelength

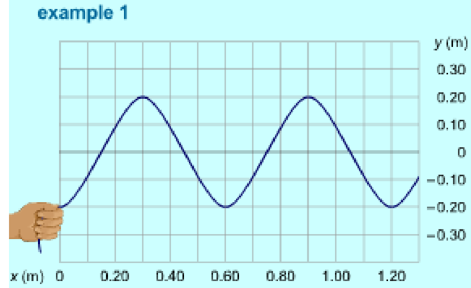
Distance between adjacent wave peaks

concept 2

	Wavelength (m)
Tsunami	~ 5.0×10^{5}
AM radio waves	~ 3.0×10^{2}
Low range whale call (100Hz)	1.5×10^{1}
4^{th} octave A (440 Hz)	7.7×10^{-1}
Bat sonar	~ 5.0×10^{-3}
X-ray waves	10^{-10}

A variety of wavelengths

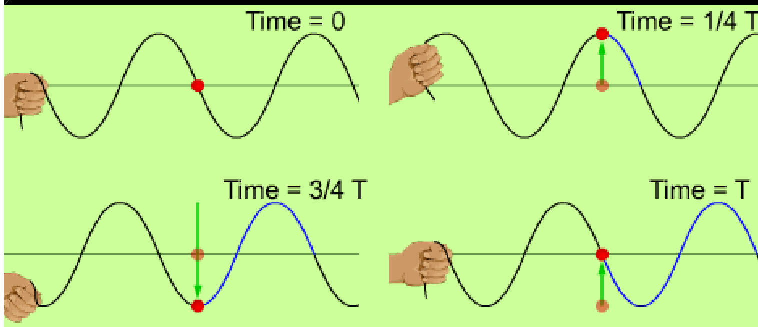
example 1



What is the wavelength of this wave?

λ = 0.60 m

16.6 - Period and frequency



concept 1

Period T of a wave
Time a particle takes to complete a cycle of motion

Period: Amount of time for a particle in a wave to complete a cycle of motion.

Frequency: Number of wave cycles per second.

The definitions of period and frequency may look familiar from your study of simple harmonic motion.

The period of a wave equals the amount of time required for a particle of the medium to move through a complete cycle of motion.

At the top of this section are four time-lapse “snapshots” of a transverse wave moving through a string. The particle marked in red moves vertically up and down. The amount of time it takes to rise to a peak, fall to a trough and return to its initial position is the period. Because the period is an interval of time, its unit is the second.

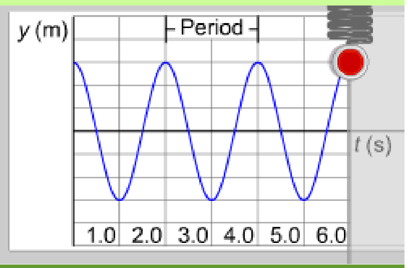
As the particle oscillates up and down through a full cycle of motion, the wave travels to the right a distance of one wavelength. Frequency is the number of full cycles of motion per second. Frequency (cycles/second) equals the reciprocal of the period (seconds/cycle). The unit for frequency is the hertz (Hz), equal to one cycle per second.

In Concept 2, we show a graph related to the transverse wave, which is not a depiction of the wave itself, but a graph of the motion of a particle over time. The scale of its horizontal axis is time, **not** position. The particle oscillates up and down in SHM, like the red particle used in the wave illustration at the top of this section. This graph could be generated in a fashion akin to the graphs you saw in the chapter on simple harmonic motion, where we rolled graph paper below a mass that had a “pen” attached to it. In this case, we would roll the paper under the red circle to record its location over time.

The period of the wave itself can be measured as the difference in time between two equivalent points (such as adjacent peaks) on the graph. The frequency of the wave is the reciprocal of the period. The particle in Concept 2 has a period of 2.0 seconds, so its frequency is 0.5 cycles per second, and that is the frequency of the associated wave.

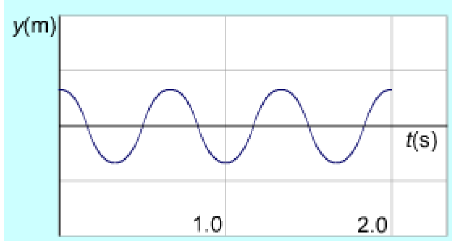
The example problem asks you to determine the frequency and period of another transverse wave. The transverse motion of a single particle in this wave is graphed with time on the horizontal axis.

concept 2



Period and frequency
Period: time to complete a cycle
· Units: seconds (s)
Frequency: number of cycles per second
· Reciprocal of period
· Units: hertz (Hz)

example 1



What are the frequency f and the period T of the wave?

$f = 3.0$ cycles/2.0 s

$f = 1.5$ Hz

$T = 1/f = 1/1.5$ cycles

$T = 0.67$ s

16.7 - Wave speed

How fast a wave moves through a medium is called its wave speed. Different types of waves have vastly different speeds, from 300,000,000 m/s for light to 343 m/s for sound in air to less than 1 m/s for a typical ocean wave. The wave in the string to the right might be moving at, say, 15 m/s.

The speed of a mechanical wave depends solely on the properties of the medium through which it travels. For example, the speed of a wave in a string depends on the linear mass density and tension of the string. This relationship is explored in another section.

For periodic waves there is an algebraic relationship between wave speed, wavelength and period. This relationship is shown in Equation 1.

This relationship can be derived by considering some of the essential properties of a wave. Wavelength is the distance between two adjacent wave peaks. The period is the time that elapses when the wave travels a distance of one wavelength. If you divide wavelength by period, you are dividing displacement by elapsed time. This is the definition of speed.

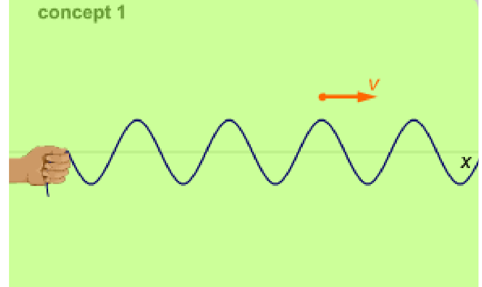
Because frequency is the reciprocal of period, the speed of a wave also equals the wavelength times its frequency. Both of these formulations are shown to the right.

Since the speed of a wave is dictated by the physical characteristics of its medium, its speed must be constant in that medium. In the example of the string mentioned above, the constant speed is determined by the linear density and tension of the string.

Because the speed of a wave in a medium is constant, the product of its wavelength and frequency is a constant. This means that for a wave in a given medium the wavelength is inversely proportional to the frequency. Increase the frequency of the wave and the wavelength decreases. Decrease the frequency and the wavelength increases.

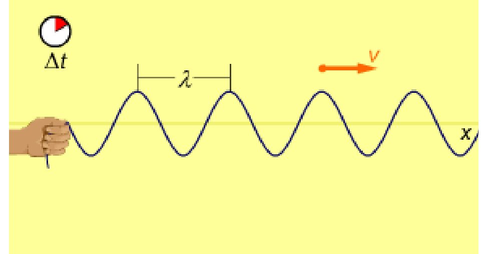
Consider sound waves. Different sounds can have different frequencies. If this were not the case, there would be no music. For example, consider the first four notes of Beethoven's Fifth Symphony (the famous "dut dut dut daaah"). As played by the violins, the first three identical notes are the G above middle C and have a frequency of 784.3 Hz and a wavelength of 0.434 m, while the fourth note (E flat) has a frequency of 622.4 Hz and wavelength of 0.547 m.

concept 1



Wave speed
How fast a wave travels

equation 1

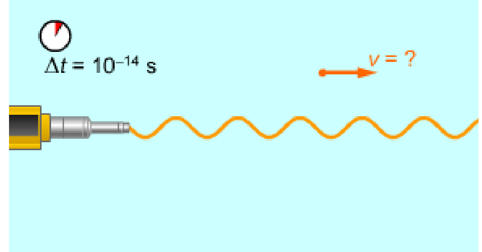


Wave speed

$$v = \frac{\lambda}{T} = \lambda f$$

v = wave speed
λ = wavelength
T = period
f = frequency

example 1



The light's wavelength is 6.0×10^{-7} m. The light completes 5.0 cycles in 10^{-14} seconds. What is the light's wave speed?

$v = \lambda f$

$$f = \frac{5.0 \text{ cycles}}{10^{-14} \text{ s}} = 5.0\times10^{14} \text{ Hz}$$

$v = (6.0\times10^{-7} \text{ m})(5.0\times10^{14} \text{ Hz})$

$v = 3.0\times10^{8}$ m/s

16.8 - Wave speed in a string

This section examines in detail the physical factors that determine the speed of a transverse wave in a string. The factors are the force on the string (the string's tension) and the string's linear density. Linear density is the mass per unit length, m/L. It is represented with the Greek letter μ (pronounced "mew").

The relationship of wave speed to the string's tension and linear density is expressed in Equation 1. The equation states that the wave speed equals the square root of the string tension divided by the linear density of the string.

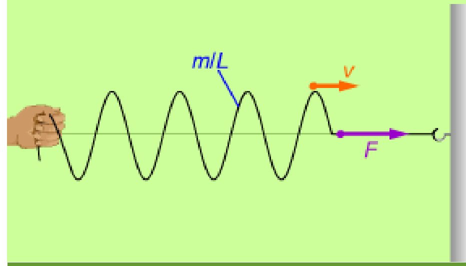
Your physics intuition may help you understand why wave speed increases with string tension and decreases with string density.

Consider Newton's second law, $F = ma$. If the mass of a particle is fixed, a larger force on the particle will result in a greater acceleration. When a string under tension is shaken up and down, the tension acts as a restoring force on the string, pulling its particles back toward their rest positions. The greater the tension, the greater this restoring force and the faster the string will return to equilibrium. This means the string will oscillate faster (its frequency increases). Because wave speed is proportional to frequency, the speed will increase with the tension.

Now let's assume that the tension is fixed, and compare wave speeds in strings that have differing linear densities (mass per unit length). Newton's second law says that for a given restoring force (tension), the particles in the more massive string will have less acceleration and move back to their rest positions more slowly. The wave frequency and wave speed will be less.

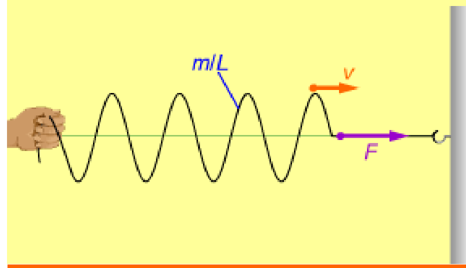
The equation on the right is a good approximation when the amplitude of the wave is significantly smaller than the overall length of the medium though which the wave moves. The force that causes the wave must also be significantly less than the tension for this equation to be accurate. The equation can be derived using the principles discussed in this section.

concept 1



Wave speed in a string
Increases with string's tension
Decreases with string's linear density

equation 1

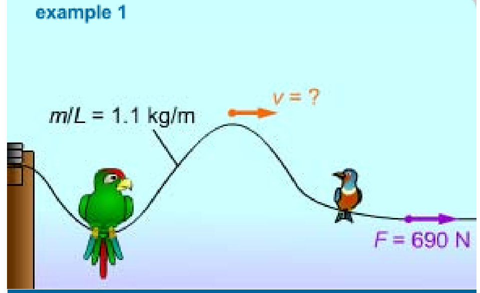


Wave speed in a string

$$v = \sqrt{\frac{F}{m/L}}$$

v = wave speed
F = string tension
m = string mass
L = string length

example 1



The green parrot is trying to dislodge the other bird. How fast will the wave he creates travel?

$$v = \sqrt{\frac{F}{m/L}}$$

$$v = \sqrt{\frac{690\ \text{N}}{1.1\ \text{kg/m}}} = 25\ \text{m/s}$$

16.9 - Interactive problem: wave speed in a string

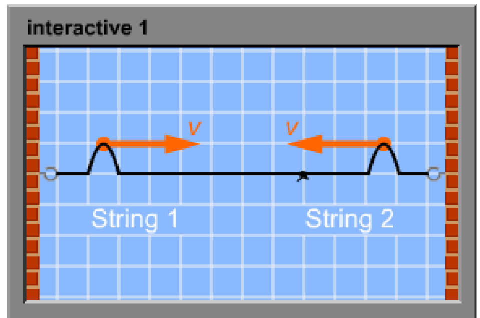
In this interactive problem, two strings are tied together with a knot and stretched between two hooks. String 1, on the left, is twice as long as string 2, on the right. Both strings have the same tension.

In the simulation, each string is plucked at its hook at the same instant. The resulting wave pulses travel inward toward the knot. The wave pulse in string 1 starts at twice the distance from the knot as the wave pulse in string 2.

You want the wave pulses to meet at the knot at the same instant. To accomplish this, set the linear density of each segment of string. When you increase a string segment's linear density in this simulation, the string gets thicker.

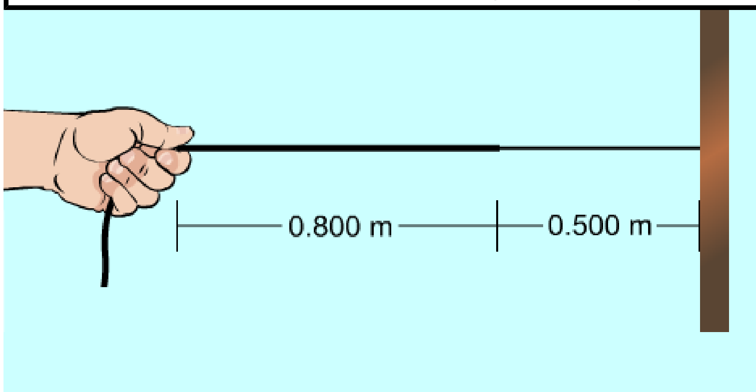
The minimum linear densities you can set are 0.010 kg/m. Set a convenient linear density for string 1; your choice for this linear density determines the appropriate linear density for string 2, which you must calculate.

Enter these values using the dials in the control panel and press GO to start the wave pulses. If they do not meet at the knot at the same instant, the simulation will pause when either of the pulses reaches the knot. Press RESET and enter different values for the linear densities to try again.



Adjust wave speed in two strings

16.10 - Interactive checkpoint: two strings spliced together



Two strings of different linear densities have been spliced together. You fix one end of the combined string to a support and hold the other in your hand. Starting from your hand, there is a 0.800 meter length of 3.40×10^{-4} kg/m density string and then a 0.500 meter length of 1.20×10^{-4} kg/m density string. You apply a tension force of 105 N. How long does it take for a pulse to get from your hand to the support?

Answer:

$t =$ ______ s

16.11 - Mathematical description of a wave

When a mechanical wave travels through a medium, the particles in the medium oscillate. Consider the diagram in Concept 1 showing a transverse periodic wave. The particles of the string oscillate vertically and the wave moves horizontally.

The vertical displacement of the highlighted particle will change over time as it oscillates. We show its displacement in Concept 1 at an instant in time.

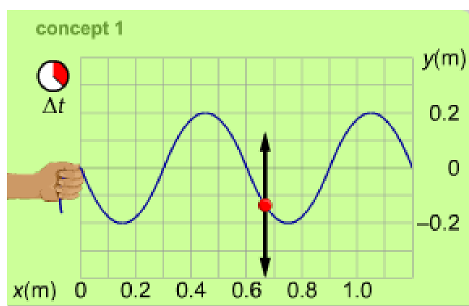
In this section, we analyze a wave in which the particles oscillate in simple harmonic motion. An equation that includes the sine function is used to describe a particle's displacement. The equation relates the vertical displacement of the particle to various factors: the horizontal position of the particle, the elapsed time, and the wave's amplitude, frequency and wavelength. When all these factors are known, the vertical position of a point can be determined at any time t.

Equation 1 describes a wave moving from left to right. The variable y in the equation is the vertical displacement of a particle at a given horizontal position away from its equilibrium position at a particular time. To use the equation, you must assume the wave has traveled the length of the string, and the time t is some time after this has occurred.

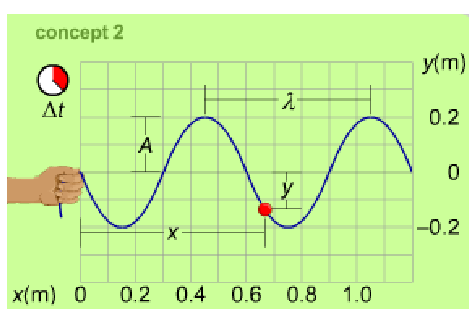
In the equation, the variable x is the particle's position along the x axis, which does not change for a given particle. The variable A is the wave's amplitude; the variable λ is the wavelength; and the variable f is the wave's frequency.

The argument of the sine function is called the *phase*. As a wave sweeps past a particle located at a horizontal position x, the phase changes linearly with respect to the elapsed time t. **The phase is an angle measured in radians**. The angle in the wave equation must be expressed in radians.

The equation in Equation 1 describes a transverse wave moving from left to right. For a wave moving from right to left, the minus sign inside the phase is switched to a plus sign, reversing the sign of the coefficient of time.



Particles
Oscillate in simple harmonic motion



Equation for traveling wave

Equation 1 assumes that a particle at position $x = 0$ at time $t = 0$ is at the equilibrium position $y = 0$. You can add what is called a phase constant to the equation to create a new equation describing a wave with a different initial state. For example, suppose a constant angle such as $\pi/2$ radians were added to the argument of the sine function. Then, the particle at $x = 0$ at time $t = 0$ would be at its maximum positive displacement, because the sine of $\pi/2$ equals one. A phase constant does not change the shape of a wave, but rather shifts it back or forward along the horizontal axis by the same amount at all times. Note that as the phase is increased by an integer multiple of 2π radians, the sine function describing the wave behaves as if there were no change at all.

If you contrast the equation here to the equation for simple harmonic motion, you will note that the equation for a traveling wave requires two inputs to determine the vertical displacement of a particle. Both equations include time, but the equation in this section also requires knowing the x position of a particle in the medium. With a wave, the vertical displacement is a function not only of time, but also of position in the medium, while the position is not a factor in SHM.

The equations to the right can be used with either transverse or longitudinal waves. When applied to longitudinal waves, the oscillation of the particles occurs parallel to the direction of travel of the wave, and then we would use the variable s instead of y to represent the horizontal displacement of a particle away from its equilibrium position.

Function relates particle's vertical displacement *y* to:
- particle's horizontal position *x*
- elapsed time *t*
- wave's amplitude, wavelength, frequency

equation 1

$$y = A \sin\left(\frac{2\pi x}{\lambda} - 2\pi f t\right)$$

Equation for traveling wave

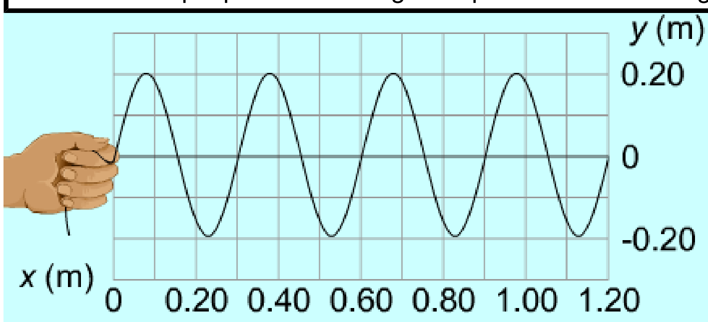
$$y = A \sin(2\pi x/\lambda - 2\pi f t)$$

y = particle's vertical displacement
A = amplitude of wave
x = particle's horizontal position
λ = wavelength, f = frequency
t = elapsed time

For wave motion toward $-x$:

$$y = A \sin(2\pi x/\lambda + 2\pi f t)$$

16.12 - Sample problem: writing an equation for a traveling wave



This is a snapshot of a traveling wave at time t = 0 s. What will be the vertical displacement of this wave at a horizontal position of 2.00 m at an elapsed time of 6.00 s? The wave's frequency is 1.33 Hz.

The illustration above shows the wave at the time $t = 0$ s and the wave is traveling left to right.

Variables

frequency	f = 1.33 Hz
wavelength	λ
amplitude	A
time	t = 6.00 s
horizontal position	x = 2.00 m
vertical displacement	y

What is the strategy?

1. Use the equation for a traveling wave that is moving to the right.
2. Analyze the diagram to determine the amplitude and wavelength.
3. Substitute all the determined quantities into the equation and evaluate the vertical displacement.

Physics principles and equations

The equation for a wave moving in the positive x direction is

$$y = A\sin\left(\frac{2\pi x}{\lambda} - 2\pi f t\right)$$

Step-by-step solution

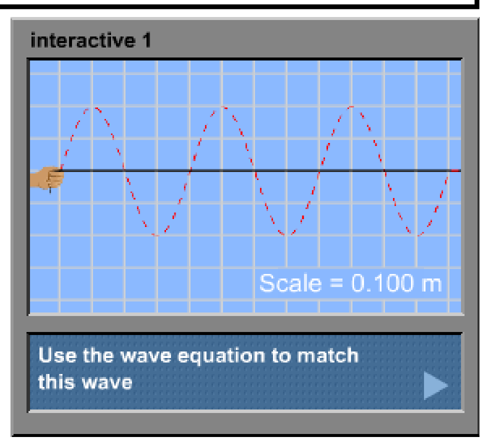
We state the equation for the wave, and substitute values that are stated in the problem or determined from the diagram.

Step	Reason
1. $y = A\sin\left(\frac{2\pi x}{\lambda} - 2\pi f t\right)$	wave equation
2. $y = (0.200\text{ m})\sin\left(\frac{2\pi x}{\lambda} - 2\pi f t\right)$	substitute amplitude
3. $y = (0.200\text{ m})\sin\left(\frac{2\pi x}{0.300\text{ m}} - 2\pi f t\right)$	substitute wavelength
4. $y = (0.200\text{ m})\sin\left(\frac{2\pi x}{0.300} - 2\pi(1.33\text{ Hz})t\right)$	substitute frequency
5. $y = (0.200\text{ m})\sin(20.9x - 8.36t)$	simplify

Now that we have determined the equation for the wave, we can solve for the vertical displacement y of the wave at a point 2.00 meters along the horizontal axis after an elapsed time of 6.00 s.

Step	Reason
6. $y = (0.200\text{ m})\sin((20.9\text{ m}^{-1})(2.00\text{ m}) - 8.36t)$	substitute horizontal position
7. $y = (0.200\text{ m})\sin(41.8 - (8.36\text{ Hz})(6.00\text{ s}))$	substitute time
8. $y = (0.200\text{ m})\sin(-8.36\text{ rad})$	simplify
9. $y = (0.200\text{ m})(-0.875) = -0.175\text{ m}$	evaluate

16.13 - Interactive problem: match the wave



The diagram on the right is a snapshot of a wave that started at the left and traveled along a stretched string for 3.00 seconds. In this interactive problem, your goal is to use the equation that describes a traveling wave to program a wave that exactly matches the target wave.

You can determine the target wave's amplitude and wavelength in the diagram to the right, or in the simulation. You can also calculate the wave speed by noting how far the target wave has traveled down the string in 3.00 seconds. With all these properties of the target wave known, you can calculate its frequency.

Using amplitude, wavelength and frequency, write an equation for the wave of the following form:

$$y = A\sin\left(\frac{2\pi}{\lambda}x - 2\pi f t\right)$$

In the control panel in the simulation, you will see this equation:

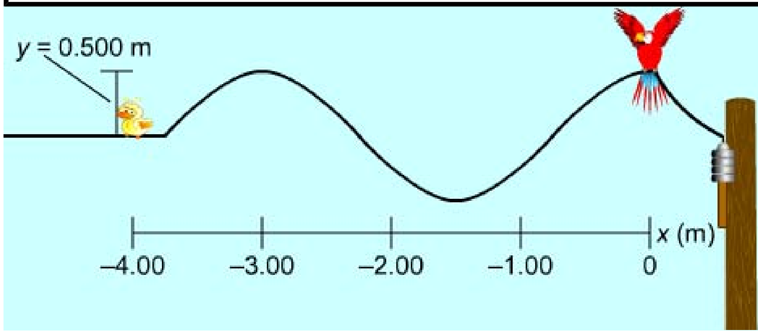
$$y = \boxed{0.100}\ \sin(\boxed{10.0}\ x - \boxed{5.00}\ t)$$

Change the values on the gauges that correspond to A, $2\pi/\lambda$ (the coefficient of x) and $2\pi f$ (the coefficient of t) to the values that you have calculated. When you are done, press GO to watch your wave propagate down the string for 3.00 seconds.

Note that if you enter coefficients corresponding to values for the wavelength and frequency that involve a change in the product λf, you are requiring that the wave speed change. The wave speed can only change if the string's tension or linear density changes. The simulation is "aware" of this fact and it adjusts the string's linear density "behind the scenes" to be compatible with the wavelength and frequency you specify.

If you do not match the target wave exactly, press RESET, enter new parameters in the equation, and press GO again to watch a new wave travel down the string.

16.14 - Sample problem: using an equation for a traveling wave



Write an equation for a wave that will displace the baby bird 0.500 m vertically 3.00 seconds after the wave starts down the wire. The wave speed is 1.50 m/s.

The red parrot on the right is jostling the wire, trying to create a wave that will displace the baby bird vertically by 0.500 m at exactly 3.00 seconds. This sample problem will help prepare you for an interactive problem in a later section.

Variables

frequency	f
wavelength	λ
amplitude	A
time	$t = 3.00$ s
horizontal position	$x = -4.00$ m
vertical displacement	$y = 0.500$ m
wave speed	$v = 1.50$ m/s

What is the strategy?

1. Use the equation for a wave moving to the left.
2. Because the baby bird must be displaced 0.500 meters above the wire's rest position, choose an amplitude for the wave that is at least 0.500 meters. We chose 0.750 m (other choices are possible).
3. Solve for the value of the argument of the sine function by dividing both sides of the equation by the amplitude and taking an arcsine. The value of the resulting angle is in radians.
4. The argument of the sine function contains two unknowns. Substitute the wave speed equation into the wavelength variable.
5. Solve for the frequency.
6. Use the wave speed equation and the frequency to solve for the wavelength.
7. Use the calculated values for the frequency and the wavelength to write the equation for the wave in terms of x and t.

Physics principles and equations

The equation for a wave moving in the negative x direction is

$$y = A \sin\left(\frac{2\pi x}{\lambda} + 2\pi f t\right)$$

The speed of a wave is related to its frequency and wavelength by

$$v = \lambda f$$

Step-by-step solution

In the first series of steps we write the equation for the wave, and then solve for and evaluate the phase.

Step	Reason
1. $y = A\sin\left(\frac{2\pi x}{\lambda} + 2\pi f t\right)$	wave equation
2. $0.500\text{ m} = A\sin\left(\frac{(2\pi)(-4.00\text{ m})}{\lambda} + (2\pi f)(3.00\text{ s})\right)$	substitute known values
3. $0.500 = 0.750\text{ m}\sin\left(\frac{(2\pi)(-4.00)}{\lambda} + (2\pi f)(3.00)\right)$	choose amplitude
4. $0.667 = \sin\left(\frac{-8.00\pi}{\lambda} + 6.00\pi f\right)$	divide and simplify
5. $\arcsin(0.667) = \left(\frac{-8.00\pi}{\lambda} + 6.00\pi f\right)$	take arcsine
6. $0.730\text{ rad} = \left(\frac{-8.00\pi}{\lambda} + 6.00\pi f\right)$	evaluate

We can solve the last equation for either wavelength or frequency by using the wave speed equation.

Step	Reason
7. $\lambda = \frac{v}{f}$	wave speed equation
8. $0.730\text{ rad} = \left(\frac{-8.00\pi}{\frac{v}{f}} + 6.00\pi f\right)$	substitute into wave equation
9. $0.730\text{ rad} = \left(\frac{-8.00\pi}{\frac{1.50\text{ m/s}}{f}} + 6.00\pi f\right)$	substitute wave speed value
10. $0.730\text{ rad} = 0.667\pi f$	simplify
11. $f = 0.348\text{ Hz}$	solve for frequency

Now that we know the wave frequency, we can calculate the wavelength.

Step	Reason
12. $\lambda = \frac{1.50\text{ m/s}}{0.348\text{ Hz}} = 4.31\text{ m}$	substitute values into equation 7 and solve

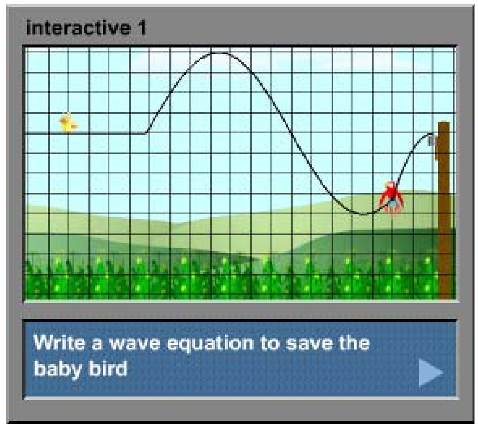
Now we can write a general wave equation describing the wave.

Step	Reason
13. $y = A\sin\left(\frac{2\pi x}{\lambda} + 2\pi f t\right)$	wave equation
14. $y = 0.750\sin\left(\frac{2\pi x}{4.31\text{ m}} + 2\pi\,(0.348\text{ Hz})t\right)$	substitute values
15. $y = 0.750\sin(1.46x + 2.19t)$	simplify

Because a wave of this type repeats itself when the argument of the sine function is increased or decreased by 2π radians, there are other correct answers to this problem. Any wave with this same amplitude (0.750 m) in which the phase is $0.730 + n2\pi$ radians (n is an integer) would also displace the baby bird by 0.500 m vertically at an elapsed time of 3.00 s.

To find additional correct answers, we also could have chosen the wavelength or frequency first, and then determined the necessary amplitude.

16.15 - Interactive problem: a well-timed wave



A baby bird is in distress, and only you can save it.

In the diagram to the right, a yellow baby bird that cannot fly is trapped on the telephone wire. Its mother is going to fly by in 4.00 seconds to try to pluck it from the wire.

To avoid coming too close to the wire, the mother bird is going to swoop in 1.00 meter above the equilibrium position of the wire.

You are the red parrot at the right end of the wire. The baby bird is at a horizontal position −4.00 meters away from you. You can see that the mother bird is going to miss the baby bird by passing one meter above it.

Then you think: If I, the noble red parrot, can create a wave in the telephone wire that will displace the fledgling one meter vertically at $t = 4.00$ seconds, then the mother bird can pluck her baby off the wire.

This, then, is your challenge: Using the equation that describes a wave traveling from right to left, specify a wave that will displace vertically by one meter a particle located at a horizontal position of −4.00 meters after exactly 4.00 seconds.

The speed of waves in this wire is 2.00 meters per second.

In the simulation control panel, you will see an equation of the following form, describing a wave that moves from right to left:

$$y = \boxed{0.10}\ \sin(\boxed{--}\ x + \boxed{3.00}\ t)$$

You can enter values that correspond to A and $2\pi f$ (the coefficient of t). The value for $2\pi/\lambda$ (the coefficient of x) will be calculated automatically depending on what you enter for $2\pi f$ (the wave speed in the telephone wire is fixed so whatever you select for the wave's frequency will determine its wavelength). Enter values and press GO to watch your wave propagate down the wire.

You will see the mother bird swoop in to a position 1.00 meter above the wire and try to pluck the baby bird off the wire at the time $t = 4.00$ s.

If you fail to help the mother rescue the baby bird, press RESET, enter new values in the equation, and press GO again to watch a new wave travel down the wire.

If you need help with this interactive problem, review the sample problem in the previous section.

16.16 - Gotchas

A mechanical wave can travel with or without a medium. No. Mechanical waves must have a medium. This is why, in the vacuum of space, there is total silence. Sound, a mechanical wave, cannot travel without a medium such as air.

The medium carrying a wave does not move along with the wave. That is correct. The medium oscillates, but it does not travel with the wave. This differentiates wind, which consists of moving air, from a sound wave. With a sound wave, the air remains in place after the sound wave has passed through.

The amplitude of a wave has no effect on the speed of the wave. That is correct. The speed of a wave is determined by the properties of the medium. This means that if you are in a hurry, it is no use yelling at people!

Wave speed in a string is a function of frequency, so if I increase the wave frequency, the wave speed will increase, too. No. The speed of a wave in a string is fixed by the tension and linear density of the string. Increasing wave frequency will cause a decrease in wavelength, but no change in wave speed.

Amplitude is the same as the vertical displacement y of a particle in a wave. No, the amplitude A is the maximum positive vertical displacement of a particle, while at a time *t* the instantaneous vertical displacement y can be anywhere between $+A$ and $-A$.

16.17 - Summary

Mechanical waves are oscillations in a medium. This chapter discussed traveling waves: disturbances that move through a medium.

There are two basic wave types. In a transverse wave, the particles in the medium oscillate perpendicularly to the direction that the wave travels. In a longitudinal wave, the particles oscillate parallel to the direction the wave travels. Some waves, such as water waves, exhibit both transverse and longitudinal oscillation.

A wave can come as a single pulse, or as a periodic (repeating) wave. In this chapter, we analyze periodic waves whose particles oscillate in simple harmonic motion.

The amplitude of a wave is the distance from its equilibrium position to its peak. The wavelength is the distance between two adjacent peaks. The period is the time it takes for a particle to complete one cycle of motion, for example, moving from peak to peak. Frequency is the number of cycles that are completed per second.

Wave speed is the speed with which a wave moves through a medium. It is equal to the wavelength of the wave times its frequency.

The wave speed in a string increases with the tension of the string and decreases with the string's linear density.

The wave equation describes the vertical displacement of a particle in a transverse wave as a function of time, the horizontal position of the particle, and the wave's frequency and wavelength.

Equations

$T = 1/f$

Wave speed

$$v = \frac{\lambda}{T} = \lambda f$$

Wave speed in a string

$$v = \sqrt{\frac{F}{m/L}} = \sqrt{\frac{F}{\mu}}$$

Wave motion toward $+x$

$$y = A \sin(2\pi x/\lambda - 2\pi f t)$$

Chapter 16 Problems

Conceptual Problems

C.1 Many children are lined up at an ice cream stand. If the child at the back pushes the child in front of him, and she in turn pushes the child in front of her, and so on, will they create a transverse or longitudinal wave in the line? Explain.

- ○ Transverse
- ○ Longitudinal

C.2 When sports spectators do "the wave," are they making a transverse or longitudinal wave? Explain.

- ○ Transverse
- ○ Longitudinal

C.3 Suppose a wave is moving along a chain at 10 m/s. Does that mean each link of the chain moves at 10 m/s? Explain.

○ Yes ○ No

C.4 A wave is traveling through a particular medium. The wave source is then modified so that it now emits waves at a higher frequency. Does the wavelength increase, decrease, stay the same, or does it depend on other factors? Explain.

i. Increases
ii. Decreases
iii. Stays the same
iv. Depends on other factors

C.5 Radio waves all travel through space at the same speed. If an AM station broadcasts at a frequency 7×10^5 Hz and an FM station broadcasts at 90.3×10^6 Hz, which station's radio waves have the longer wavelength? Explain.

i. AM station
ii. FM station
iii. Both the same

C.6 You have two strings of different linear densities tied together and stretched end-to-end between a pair of supports. What happens to the speed of a wave that passes from the more dense string to the less dense one?

i. The speed decreases
ii. The speed increases
iii. No change

C.7 Waves of identical frequency and amplitude are simultaneously launched down two identical wires of equal mass and length, although one wire is tauter than the other. Which wave will arrive at the other end first, or will they arrive at the same time?

i. Wave on tauter string arrives first.
ii. Wave on looser string arrives first.
iii. They arrive at the same time.

C.8 Earthquake waves in the substance of our planet occur in two types: P waves, which are longitudinal waves, and S waves, which are transverse waves. In seismology, P and S usually stand for primary and secondary. Explain how either of the following could also be valid associations of the letters P and S. (a) P is for push-pull waves, S is for sideways waves. (b) P is for pressure waves, S is for shearing waves.

(a) Submit answer on paper
(b) Submit answer on paper

Section Problems

Section 0 - Introduction

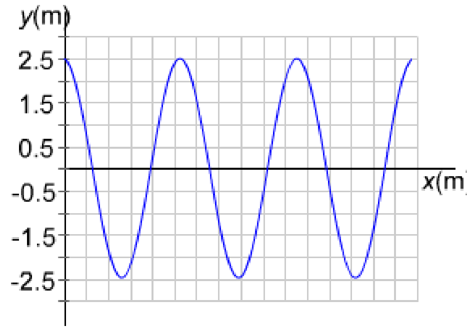
0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Does changing the frequency of a wave in the simulation change its wavelength? (b) Does changing the frequency of a wave in the simulation change its amplitude? (c) Does changing the amplitude in the simulation cause the the wave to move noticeably faster or slower down the string?

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c) ○ Yes ○ No

Section 4 - Amplitude

4.1 What is the amplitude of the wave that is shown? Enter the answer using two significant figures.

__________ m



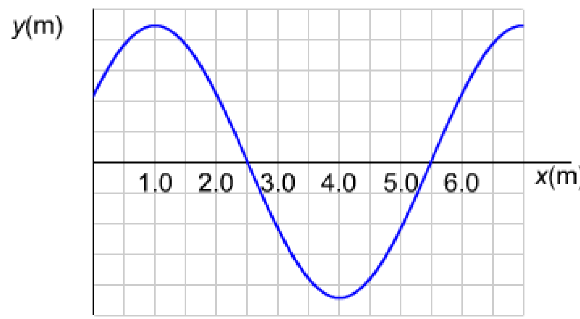
4.2 A water strider sits on a the surface of a still lake. A rock thrown into the lake creates a series of sinusoidal ripples that pass through the location where the strider sits, so that it rises and falls. The distance between its highest and lowest locations is 0.008 m. What is the amplitude of the wave?

i. 0.002 m
ii. 0.004 m
iii. 0.008 m

Section 5 - Wavelength

5.1 What is the wavelength of the wave that is shown? Enter the answer using two significant figures.

__________ m



5.2 In a foggy harbor, a tugboat sounds its foghorn. Bobbie stands on shore, 7.5×10^2 m away. The foghorn's sound wave completes 1.1×10^3 cycles on its way to Bobbie. What is the wavelength of the sound wave?

__________ m

Section 6 - Period and frequency

6.1 A wave has a frequency of 575 Hz. What is its period?

__________ s

6.2 If 12 wave crests pass you in 26 seconds as you bob in the ocean, what is the frequency of the waves?

__________ Hz

Section 7 - Wave speed

7.1 An important wavelength of radiation used in radio astronomy is 21.1 cm. (This wavelength of radiation is emitted by excited neutral hydrogen atoms.) This radiation travels at the speed of light, 3.00×10^8 m/s. Compute the frequency of this radio wave.

__________ Hz

7.2 A wave has a speed of 351 m/s and a wavelength of 2.4000023468268647 meters. What is its period?

__________ s

7.3 Suppose you are standing on a boat in the ocean looking out at a person swimming in the water. The waves swell under your boat at a frequency of 0.80 Hz, while traveling with a wave speed of 1.5 meters per second. At any given moment, you count that there are exactly 6 complete waves between you and the swimmer. How far away is the swimmer?

__________ meters

7.4 The radio wave from a station broadcasting at 1.3×10^6 Hz has a wavelength of 230 m. What is the wave's speed?

__________ m/s

7.5 Two waves travel in the same medium at the same speed. One wave has frequency 5.72×10^5 Hz and wavelength 0.533 m. The other wave has frequency 6.13×10^5 Hz. What is the second wave's wavelength?

__________ m

7.6 Two waves travel in the same medium at the same speed. One has wavelength 0.0382 m and frequency 9.67×10^6 Hz. The other has wavelength 0.0446000398122414 m. What is the period of the second wave?

__________ s

Section 8 - Wave speed in a string

8.1 A wave is traveling through a 35.0-meter-long cable strung with a tension of 35,000 newtons. The mass of this length of cable is 10.2 kilograms. What is the speed of a wave that is traveling in the cable?

__________ m/s

8.2 A 0.340 gram wire is stretched between 2 points that are 75.0 cm apart. The tension in the wire is 620 N. When the string is plucked, a wave is created with wavelength equal to twice the length of the string. This wave's frequency is called the "fundamental frequency" of the string. What is this fundamental frequency?

__________ Hz

8.3 The speed of longitudinal waves in a fluid (sound waves, for example) is given by

$$v = \sqrt{\frac{B}{\rho}}$$

where B is a constant property of the fluid called its bulk modulus and ρ is the density of the fluid. For water at 4ºC, $B = 2.15\times10^9\ \mathrm{N/m^2}$ and $\rho = 1000\ \mathrm{kg/m^3}$. What is the speed of longitudinal waves in water at this temperature?

__________ m/s

8.4 When a certain string has a tension of 34 N, a wave travels along it at 68 m/s. What is the linear density of the string?

__________ kg/m

Section 9 - Interactive problem: wave speed in a string

9.1 Use the information given in the interactive problem in this section to answer the following question. If the linear density of string 1 is 0.020 kg/m, what linear density for string 2 will cause the wave pulses will reach the knot at the same instant? Test your answser using the simulation.

__________ kg/m

Section 10 - Interactive checkpoint: two strings spliced together

10.1 A single string is made up of two strings of different densities joined together. Each of the strings is 1.40 m long. The first string has linear density 2.30×10^{-4} kg/m and the second has linear density 1.80×10^{-4} kg/m. If it takes 2.20×10^{-3} s for a wave to travel from the beginning to the end of the combined string, what is the tension on the combined string?

__________ N

10.2 A single string of length 3.22 m is made up of two separate shorter strings, A and B, spliced together. String A has linear density 3.12×10^{-4} kg/m and string B has linear density 1.87×10^{-4} kg/m. The tension on the combined string is 231 N. It takes a wave 3.15×10^{-3} s to travel from the beginning to the end of the combined string. What is the length of string A?

__________ m

Section 11 - Mathematical description of a wave

11.1 A wave is described by the equation $y = (4.5\times10^{-2}) \sin (1.9x - 3.2t)$, where lengths are measured in meters and time in seconds. What is the displacement of a particle at position $x = 2.7$ m at time $t = 5.8$ s? Note: Remember that the argument of the sine function in the wave equation is expressed in radians.

__________ m

11.2 A wave is defined by the equation $y = 2.1 \sin(4.2\pi x + 6.4\pi t)$. Time is measured in seconds and lengths are in meters. What are the wave's (a) direction of motion; (b) amplitude; (c) wavelength; (d) frequency?

(a) i. Left to Right
ii. Right to Left
iii. Up to down
iv. Down to Up

(b) __________ m
(c) __________ m
(d) __________ Hz

11.3 A taut string is agitated by a mechanical oscillator, producing a transverse wave that is described by the equation

$$y = (3.34 \times 10^{-3}) \sin(1.57x + 94.2t + 0.222)$$

where time is measured in seconds, and lengths in meters. Find (a) the direction of travel of the wave, (b) the wave's amplitude, (c) the wavelength, and (d) the frequency.

(a) i. Left to right
ii. Right to left
iii. Up to down
iv. Down to up

(b) __________ m
(c) __________ m
(d) __________ Hz

11.4 A wave is described by the equation $y = A \sin(2.3x - 1.8t)$, where lengths are measured in meters, and time in seconds. What is the earliest time t (> 0) when the displacement of a particle at position $x = 1.6$ m is at a maximum?

__________ s

Section 13 - Interactive problem: match the wave

13.1 Use the information given in the interactive problem in this section to calculate the (a) amplitude, (b) angular wave number and (c) angular frequency required to match the wave. Test your answer using the simulation.

(a) __________ m
(b) __________ rad/m
(c) __________ rad/s

Section 15 - Interactive problem: a well-timed wave

15.1 Use the information given in the interactive problem in this section to answer the following question. If the amplitude of the wave is 1.00 m, what is the smallest possible angular frequency for the wave that saves the baby bird? Test your answer using the simulation.

__________ rad/s

Additional Problems

A.1 As you stand on a pier looking at a harbor, the waves created by a large freighter come ashore. You estimate the wavelength to be 3.0 m and you count twelve wave peaks arriving in one minute. (a) What is the frequency of the wave in hertz? (b) How fast is the wave traveling?

(a) __________ Hz
(b) __________ m/s

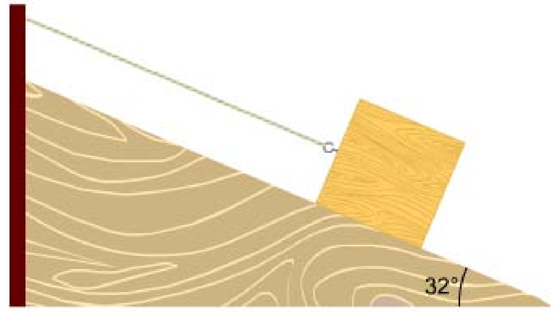
A.2 A steel cable is stretched across a gorge. The cable is 190 m long and has a mass of 44 kg. When one end of the cable is struck with a hammer, a wave pulse travels down the cable, reflects at the far end, and is detected back at the first end 14 s later. What is the speed of the wave?

__________ m/s

A.3 A steel cable suspends a grand piano from an apartment balcony, over a busy street. (It is either moving day or an avant garde art performance.) If the cable has a length of 6.8 m and a mass of 21 kg, and the piano has mass of 360 kg, how long will it take for a wave to travel the length of the cable? Note: Since the cable is so much less massive than the piano, the mass of the cable can be ignored when calculating its tension.

__________ s

A.4 A block of mass 3.1 kg rests on a frictionless inclined plane, held in place by a string of mass 0.023 kg. The angle of the plane from the horizontal is 32°, and the string length is 1.5 m. How long does it take a wave to travel from one end of the string to the other? Note: Since the string is so much less massive than the block, the mass of the string can be neglected when calculating its tension.



__________ s

A.5 A block of mass 19.3 kg hangs from an elastic string of mass 0.0226 kg. When the string is unstretched, its length is 0.760 m. The spring constant for the elastic is 88.2 N/m. The block is hanging from the string and is not moving. Note: Since the string is so much less massive than the block, the mass of the string can be ignored when you calculate its tension and extension. (a) What is the magnitude of the tension in the string in this stretched position? (b) What is the length of the string with the block attached? (c) With the block attached, what is the speed of a wave on the stretched string?

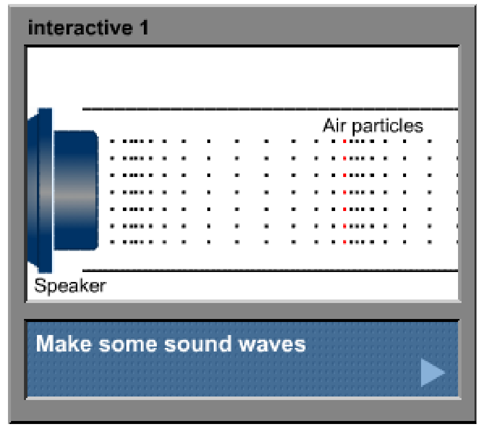
(a) __________ N

(b) __________ m

(c) __________ m/s

chapter 17 Sound

17.0 - Introduction



Sounds are so commonplace that it is easy to take them for granted, but they are a central part of the human experience. When you think of sound, you may think of your favorite song or an alarm clock that goes off early and loud. To a physicist, though, both a pleasant song and a shrill alarm are mechanical longitudinal waves consisting of regions of high and low pressure. The physics of sound waves is the topic of this chapter.

Sounds can be classified as audible, infrasonic or ultrasonic. Audible sounds are in the frequency range that can be heard by humans. *Infrasonic* sounds are at frequencies too low to be heard by humans, but animals such as elephants and whales use them to communicate over great distances. *Ultrasonic* sounds are at frequencies too high to be perceived by humans. They are used by bats for sonar and by doctors to see inside the human body.

You may have used the speed of sound in air to estimate the distance to a thunderstorm. The flash from the lightning reaches you almost instantaneously, while the sound from the thunder takes more time. Sound travels at approximately 343 m/s in air at $20°\text{C}$, so for every third of a kilometer of distance to the lightning, the sound of the thunder lags the flash of light by about one second.

Sound travels slowly enough in air that manmade objects such as airplanes can catch and pass their own sound waves. You may have heard the result when a plane is flying faster than the speed of sound: a sonic boom. A small sonic boom is also the cause of the "crack" of a whip, as the tip of the whip travels faster than the speed of sound.

You may begin your study of sound with the simulation to the right, which allows you to experiment with a loudspeaker that causes sound waves to travel through a tube filled with air particles. One set of particles is colored red to emphasize that all the particles just oscillate back and forth; they do not travel along with the wave.

Sound waves can be described with the same parameters that are used to describe transverse mechanical waves: amplitude, frequency and wavelength. Recognizing these parameters in a longitudinal wave may require some practice.

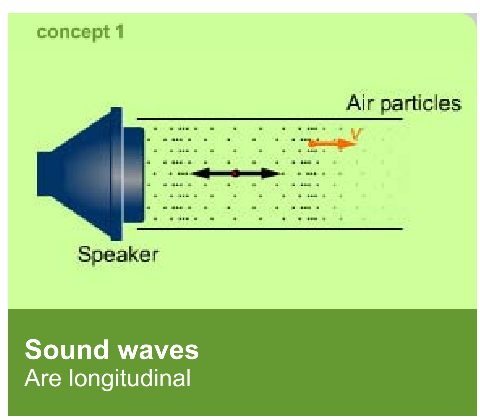
When you open the simulation, press GO to send a sound wave through the air. You will see the loudspeaker's diaphragm vibrate horizontally. This causes the nearby air particles to vibrate and a longitudinal wave to travel from left to right along the length of the tube.

Observe the differences between this wave and the transverse waves you saw in strings. You should be able to see how the particles of the medium (air) oscillate **parallel** to the direction the wave travels in a longitudinal wave, as opposed to the perpendicular motion of the particles in a transverse wave.

The simulation lets you control the loudspeaker to determine the amplitude and frequency of the wave. As with any wave, the amplitude is the maximum displacement of a particle from its rest position, and the frequency is the number of cycles per second. You can vary these parameters and observe changes in the motion of the loudspeaker and in the properties of the sound wave. You can also observe how the wavelength changes when you alter the frequency.

Humans can identify different sound waves by pitch, which is related to frequency. If you have audio on your computer, turn it on and listen to the pitch created by a particular wave. Then increase the frequency and hear how the pitch changes. The loudness of a sound wave is related to its amplitude. Increase the amplitude of the wave in the simulation and note what you hear.

17.1 - Sound waves



Sound waves are longitudinal mechanical waves in a medium like air generated by vibrations such as the plucking of a guitar string or the oscillations of a loudspeaker.

Sound waves are caused by alternating compression and decompression of a medium. In Concept 1, a loudspeaker is shown. As the loudspeaker diaphragm moves forward, it compresses the air in front of it, causing the air particles there to be closer together. This region of compressed air is called a *condensation*. The pressure and density of particles is greater in a region of condensation. This compressed region travels away from the loudspeaker at the speed of sound in air.

The diaphragm then pulls back, creating a region in which there are fewer particles. This region is a *rarefaction*, and the pressure there is lower. The rarefaction also travels away from the loudspeaker at the speed of sound. The velocity of the wave is indicated with the orange vector $\mathbf{v}$ in the diagram. The back-and-forth motion of an individual particle is indicated with the black arrows.

The illustration for Concept 2 also shows the alternating regions of condensation and rarefaction as the loudspeaker oscillates back and forth. The wavelength is the distance between two successive areas of maximum condensation or rarefaction. As with transverse waves, the wavelength is measured along the direction of travel. The wavelength can be readily visualized as the distance between the midpoints of the two regions of condensation shown in the diagram.

With a wave in a string, a complete cycle of motion occurs when a particle in the string starts at a point (perhaps a peak), moves to a trough, and returns to a peak. With a sound wave, an analogous cycle passes from compression to rarefaction and then returns to compression.

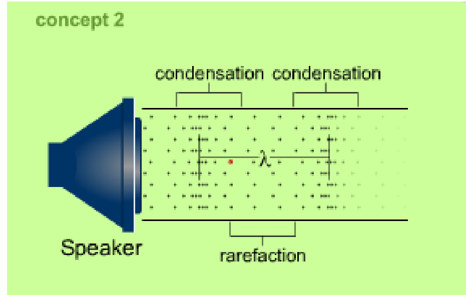
To measure the period, you can note how long it takes for an air particle to pass through a complete cycle at a given location. As with transverse waves, the frequency of a sound wave is the number of cycles completed per second.

Concept 3 illustrates the motion of an individual particle in a sound wave. Refresh the browser page to see an animation of the particle's motion. The particle oscillates back and forth horizontally as regions of high and low pressure pass by, also horizontally. The particle is first pushed to the right as an area of higher pressure passes, and then pulled to the left by a region of lower pressure. As high and low pressure regions pass by, the individual particles oscillate in simple harmonic motion.

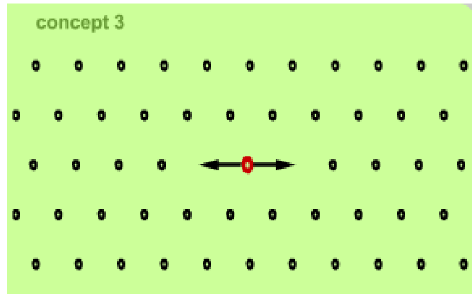
The harmonic oscillation of the particles distinguishes a sound wave from wind. Wind causes particles to have net displacement; it moves them from one location to another. There is no net movement of air particles over time due to sound waves. The particles oscillate back and forth around their original locations.

Two aspects of the behavior of gas molecules will help you understand sound waves. First, we have used increased density and increased pressure to define condensation, and decreased density and decreased pressure to define rarefaction. Density and pressure are correlated. The ideal gas law (which you may study in a later chapter) states that, everything else being equal, pressure increases with the number of gas molecules in a system. This means that the pressure is greater in regions of condensation than in regions of rarefaction. Other factors (such as temperature) also influence pressure, but in this discussion we treat them as constant.

Second, the speed of sound can be understood in terms of the behavior of air molecules. The speed v of a sound wave does not equal the speed of the loudspeaker's motion. That may seem odd. How could the waves move faster or slower than the object that causes them?



Structure of a sound wave
Condensation, followed by rarefaction



Particles in a sound wave
Move in simple harmonic motion

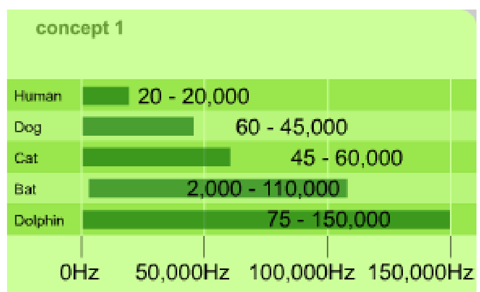
The diagram we use simplifies the nature of the motion of air molecules in a wave. Air molecules at room temperature move at high speeds (hundreds of meters per second) and frequently collide. On average, their location is stationary since their motion is random, so we can draw them as stationary to reflect their average position. On the other hand, they are always moving, and when the loudspeaker moves, it changes the velocity of molecules that are already in motion. The change in velocity caused by the loudspeaker is transmitted through the air by multiple collisions as a function of the random speeds of the molecules. The faster the molecules are moving, the more frequently they will collide. As the air becomes warmer, for example, the average thermal speed of the molecules increases, as does the speed of sound. Properties of the medium itself, not the speed of the loudspeaker, determine the speed of sound.

17.2 - Human perception of sound frequency

When a sound wave composed of alternating high and low pressure regions reaches a human ear, the wave vibrates the eardrum, a thin membrane in the outer ear. The vibrations are then carried through a series of structures to generate signals that are transmitted by the auditory nerve to the brain, which interprets them as sound.

There is a subjective relationship between the frequency of a sound wave and the *pitch* of the sound you hear. Pitch is the distinctive quality of the sound that determines whether it sounds relatively high or low within a range of musical notes. A home smoke alarm issues a high-pitched beep, while a foghorn emits a low-pitched rumble. The human ear is extremely sensitive to differences in the frequency of sound waves.

A *pure tone* sound consists of a sound of a single frequency. The tones produced by musical instruments combine waves with several frequencies, but each note has a fundamental frequency that predominates. The note middle C on a piano has a fundamental frequency of 262 cycles per second (Hz); the lowest and highest notes of a piano have frequencies of 27.5 and 4186 Hz, respectively. Orchestras tune to a note of 440 Hz (the A above middle C). To experiment with frequency and pitch yourself, try the interactive simulation in the next section.



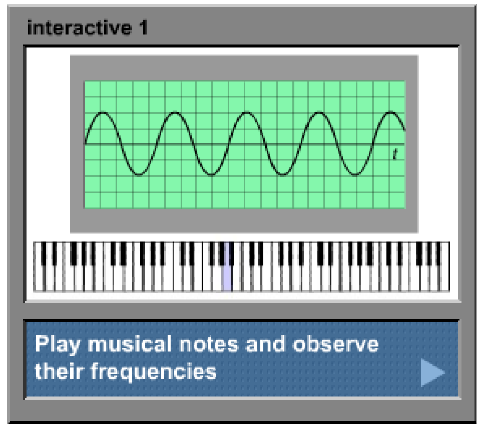
Audible frequencies

A young person can hear sounds that range from 20 to 20,000 Hz. With age, the ability of humans to perceive higher frequency sounds diminishes. Middle-aged people can hear sounds with a maximum frequency of about 14,000 Hz.

The human ear is sensitive to a wide range of frequencies, but other animals can perceive frequencies that humans cannot. Whales emit and hear sounds with a frequency as low as 15 Hz. Bats emit sounds in a frequency range from 20,000 Hz up to 100,000 Hz and then listen to the reflected sound to locate their prey. Dogs (and cats) can detect frequencies more than twice as high as humans can hear, and some dog whistles operate at frequencies that the animals can hear but humans cannot.

Interestingly, when it comes to certain sounds, the ear is not the most sensitive part of the human body. Sometimes you can feel sounds even when you cannot hear them. Some contrabass musical instruments are designed to play notes below the lowest limit of human hearing (20 Hz). For example, the organ in the Sydney (Australia) Town Hall can play a low, low C that vibrates at only 8 Hz, a rumbling that can only be felt by the audience members. Low frequency sounds are also used in movies and some arcade games to increase tension and suspense.

17.3 - Interactive problem: sound frequency



In this interactive simulation, you can experience the relationship between sound wave frequency and pitch.

The simulation includes a virtual keyboard. (As you might expect, your computer must have a sound card and speakers or a headphone for you to be able to hear the musical notes.)

Above the keyboard is an oscilloscope, used to display the sound wave. The oscilloscope graphs the waves with time on the horizontal axis. Each division on the horizontal axis represents one millisecond.

On the vertical axis, the oscilloscope graphs an air particle's displacement from equilibrium as a function of time, as the sound wave passes by. The wave is longitudinal, and peaks and troughs on the graph correspond to the particle's maximum displacement, which occurs along the direction of the wave's motion. Although we do not provide units on the vertical axis, displacements of the particles in audible sound waves are generally in the micrometer range.

The frequency of ordinary musical notes is high enough that the oscilloscope graphs time in milliseconds. The middle C on the keyboard (a white key near its midpoint) has a frequency of about 262 Hz (cycles per second), or 0.262 cycles per millisecond. A full cycle of this sound wave would span slightly less than four squares on the horizontal axis of the oscilloscope.

When you start the simulation, the instrument is set to "synthesizer." When you press a key, you will hear a tone that plays at the same intensity for as long as you hold down the key. Even this simple synthesizer tone consists of several frequencies, but one fundamental frequency predominates, and it is displayed on the oscilloscope.

You can use the oscilloscope to compare the frequencies of various sound waves, as well as using your ears to compare various pitches. Do they relate? Specifically, do higher pitched musical notes have a higher or lower frequency than lower pitched musical notes? If you know how to play notes on the piano that are an octave apart, compare the frequencies and wavelengths of these sounds. What are the relationships?

You can also set the simulation to hear notes from a grand piano. These tones are even more complex than the synthesizer and again we display just the fundamental frequency. To simulate a piano's sound, the notes will fade away even if you hold down the key, but the oscilloscope will continue to display the initial sound wave.

This simulation is designed to give you an intuitive sense of the frequencies of different musical notes. If you know how to play the piano, even something as simple as "Chopsticks," play a song and observe the waves that make up that tune. You can also play a note and see how close a friend can come to guessing it. Some people have a capability called *perfect pitch*, and can tell the note or frequency correctly every time.

17.4 - Speed of sound in various media

concept 1

Speed of sound in various media

	Substance (at 20°C)	Speed(m/s)
Gases	Air	343
	Helium (at 0°C)	965
	Hydrogen (at 27°C)	1,310
Liquids	Ethanol	1,162
	Fresh water	1,483
	Salt water (3.5% salinity)	1,522
Solids	Copper	5,010
	Glass	5,640
	Steel	5,940

Speed of sound
Depends on medium

As with other mechanical waves, the speed of a sound wave depends on the properties of the medium through which it travels. Sound moves about five times faster in water than in air and even faster in glass or steel. The table to the right shows the speed of sound in various media. In solids, sound vibrations can be transverse as well as longitudinal, as in earthquakes. The table gives values for longitudinal waves.

At the right, we also show equations for calculating the speed of sound in air, in fluids and in solids. Equation 1 shows how to calculate the speed of sound in air. This speed is a function of the random speeds of air molecules, which in turn is a function of air temperature. The molecules will move faster at higher temperatures. The faster the molecules move, the faster they can transmit a sound wave. The equation uses the speed of sound at $0°C$, 331 m/s, as a reference. The equation provides a good approximation for a wide range of temperatures.

Equation 2 shows how to calculate the speed of sound in any fluid (a liquid or a gas). The upper variable is the bulk modulus, a measure of how difficult it is to compress the fluid. Specifically, the bulk modulus is a measure of how much pressure it takes to cause a given percent change in the volume of a fluid. A material with a greater bulk modulus is harder to compress.

The variable ρ in Equation 2 is the material's density. Sound travels faster in a fluid with greater bulk modulus or lesser density.

You can use Equation 2 to compare the speed of sound in air and water. Water is much denser than air, yet sound travels considerably faster in water. For this to be true, water must have a significantly greater bulk modulus than air, and it does. It takes almost 15,000 times as much force to compress a volume of water a given fractional amount as it does to compress the same volume of air the same amount.

Equation 3 shows how to calculate the speed of a sound wave traveling in a solid. Instead of the bulk modulus, the equation uses Young's modulus, which relates stress and strain in a material. Again, ρ is the solid's density.

Equations 2 and 3 share a common feature: An elastic property of the material is in the numerator and the density of the material is in the denominator. The upper factor quantifies a material's resistance to compression. The more a material resists compression, the faster sound waves travel through it. As for the denominator, lower density means less mass per unit volume, and less mass means greater acceleration for a given force. Consequently, wave speed increases as density decreases.

You may note that these equations bear a resemblance in form to the equation for the speed of a wave in stretched string, discussed in another chapter. That equation also involves the square root of a fraction, with tension in the numerator, and a density property in the denominator.

equation 1

Speed of sound in air

$$v = (331\ \text{m/s})\sqrt{1 + \frac{T_C}{273\,^\circ\text{C}}}$$

v = wave speed
331 m/s = speed of sound in air at 0°C
T_C = temperature of air, in °C

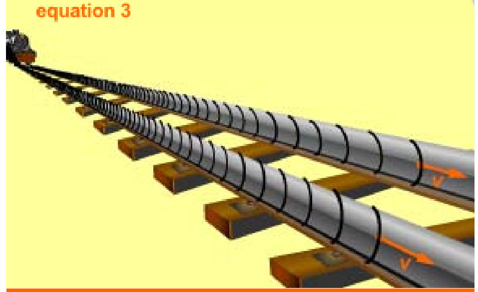
equation 2

Speed of sound in fluids

$$v = \sqrt{\frac{B}{\rho}}$$

v = wave speed
B = bulk modulus
ρ = density

equation 3



Speed of sound in solids

$$v = \sqrt{\frac{Y}{\rho}}$$

v = wave speed
Y = Young's modulus
ρ = density

17.5 - Sample problem: speed of sound

On a hot summer day (temperature 30.0°C), a hiker shouts into a canyon. She hears the echo 1.50 s later. How far away is the opposite side of the canyon?

Variables

temperature	$T = 30.0°\text{C}$
time	$\Delta t = 1.50$ s
speed of sound	v
distance to other side of canyon	d

What is the strategy?

1. Determine the speed of sound in air at 30.0°C.
2. Use the speed of sound and the time for the sound to travel out and back to determine the distance to the other side of the canyon.

Physics principles and equations

The equation for the speed of sound in air

$$v = (331 \text{ m/s})\sqrt{1+\frac{T_C}{273\,°\text{C}}}$$

The definition of speed

$$v = \frac{\Delta x}{\Delta t}$$

Step-by-step solution

First, we calculate the speed of sound on this warm day.

Step	Reason
1. $v = (331 \text{ m/s})\sqrt{1+\frac{T_C}{273\,°\text{C}}}$	speed of sound
2. $v = (331 \text{ m/s})\sqrt{1+\frac{30.0\,°\text{C}}{273\,°\text{C}}}$	substitute value
3. $v = 349$ m/s	evaluate

Now that we have determined the speed of sound, we can determine the distance to the other side of the canyon.

Step	Reason
4. $v = \frac{\Delta x}{\Delta t}$	definition of speed
5. $v = \frac{2d}{\Delta t}$	total distance
6. $d = \frac{v\Delta t}{2}$	solve for d
7. $d = \frac{(349 \text{ m/s})(1.50 \text{ s})}{2}$	substitute values
8. $d = 262$ m	evaluate

17.6 - Interactive checkpoint: speed of sound

Engineers are blasting to make way for a highway. Part of the blast area contains a long, straight section of unused railroad track. You are 2.50 kilometers away, and put your ear on the track to listen for the blast. You hear the blast transmitted through the tracks, and then at some later time you hear it through the air with your other ear. What is the time difference? The tracks are made of a steel that has a Young's modulus of 197×10^9 N/m^2 and a density of 7750 kg/m^3. The air temperature is 17.0°C.

Answer:

Δt = [ ] s

17.7 - Mathematical description of a sound wave

To the right is the equation for the displacement s of an air particle relative to its equilibrium, or rest, position as it oscillates in a sound wave. The equation resembles that of a particle in a transverse wave, but here the displacement of the particle is horizontal, along the direction of motion of the longitudinal wave.

As with a transverse wave, the displacement of a particle is a function of variables such as the time and the equilibrium position and amplitude of the particle, and constant factors that describe the sound wave, such as its wavelength and frequency.

In the equation, the variable x is the particle's horizontal equilibrium position. As the particle oscillates, s is the displacement of the particle from its equilibrium position at time t.

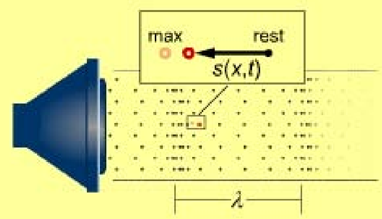
To illustrate some of the factors in the equation, we show an enlarged view of the particle in the wave. It is shown moving to the left from its rest position. The absolute value of its maximum displacement from the rest position is its amplitude. The function calculates the particle's displacement s from its rest position.

The formula in Equation 2 shows how to calculate the **change in air pressure** caused by a sound wave. This refers to the change from the undisturbed air pressure when no sound wave is present.

The air pressure in a wave varies sinusoidally over time, as well. Note that the first equation uses a cosine function and the second, a sine function. The functions are $90°$ out of phase. This means the change in pressure is at its maximum magnitude when the particle is at its equilibrium position. Conversely, the pressure change is zero when the particle is at the extremes of its motion.

Equation 3 shows how to calculate the maximum change in pressure. This factor is used in Equation 2.

equation 1

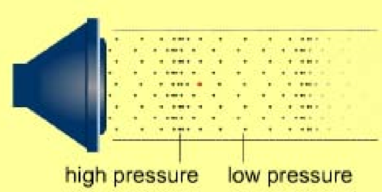


Displacement

$$s(x,t) = A\cos\left(\frac{2\pi x}{\lambda} - 2\pi f t\right)$$

s = particle's displacement
x = particle's position at equilibrium
t = time, A = particle's amplitude
λ = wavelength, f = frequency

equation 2



Air pressure

$$\Delta P(x,t) = \Delta P_{max} \sin\left(\frac{2\pi x}{\lambda} - 2\pi f t\right)$$

ΔP = change in air pressure
ΔP_{max} = pressure amplitude
(maximum pressure change)

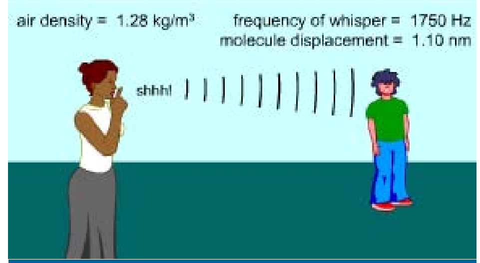
equation 3

Maximum pressure change

$$\Delta P_{max} = 2\pi A f \rho v$$

ρ = density
v = wave speed

example 1



What is the maximum amount of pressure change caused by this whisper?

$$\Delta P_{max} = 2\pi A f \rho v$$

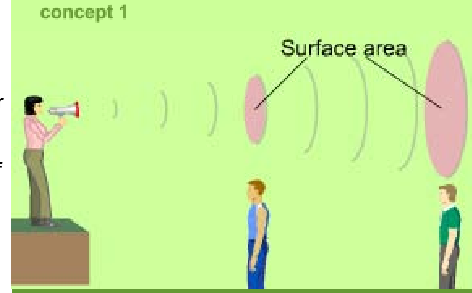
$$\Delta P_{max} = 2\pi\,(1.10 \times 10^{-9})(1750)(1.28)(343)$$

$$\Delta P_{max} = 0.00531 \text{ Pa}$$

17.8 - Sound intensity

Sound intensity: The sound power per unit area.

concept 1



Sound intensity
Sound power that passes perpendicularly through a surface area
Diminishes with square of distance from sound source

Sound carries energy. It may be a small amount, as when someone whispers in your ear, or it may be much more, as when the sonic boom of an airplane rattles windows, or when a guitar amplifier goes to 11.

Sound intensity is used to characterize the power of sound. It is defined as the power of the sound passing perpendicularly through a surface area. Watts per square meter are typical units for sound intensity.

The definition of sound intensity is shown in Equation 1. An intensity of approximately 1×10^{-12} W/m^2 is the minimum perceptible by the human ear. An intensity greater than 1 W/m^2 can damage the ear.

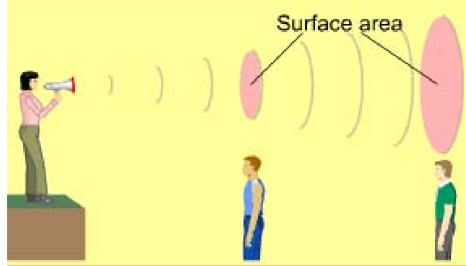
As a sound wave travels, it typically spreads out. You perceive the loudness of the sound of a loudspeaker at an outdoor concert differently at a distance of one meter than you do at 100 meters. The intensity of the sound diminishes with distance.

Equation 2 is used to calculate the intensity of sound when it spreads freely from a single source. The intensity diminishes with the square of the distance from the source. It does so because the sound energy in this case is treated as being distributed over the surface of a sphere whose radius increases with time. The denominator of the expression for intensity is $4\pi r^2$, the expression for the surface area of a sphere.

The example problem to the right asks you to find the relative sound intensities experienced by two listeners, one twice as far as the other from the fireworks. In this scenario, the sound is four times as intense for the closer listener, but this does not mean he hears the sound as four times louder. The loudness of sounds as perceived by human beings has a logarithmic relationship to sound intensity. This topic is explored in another section.

Here we have focused on sound that freely expands in all directions. However, sound can also reflect off surfaces such as walls. Concert halls are designed to take advantage of this reflection to deliver a full, rich sound to the audience.

equation 1

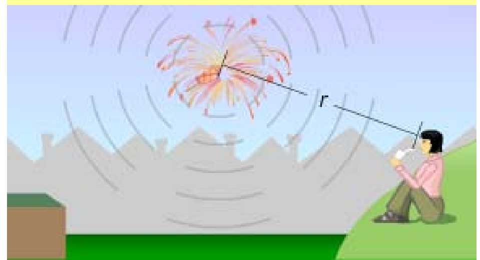


Sound intensity

$$I = P/A$$

I = sound intensity
P = power perpendicular to surface
A = surface area
Units: watts/meter^2

equation 2

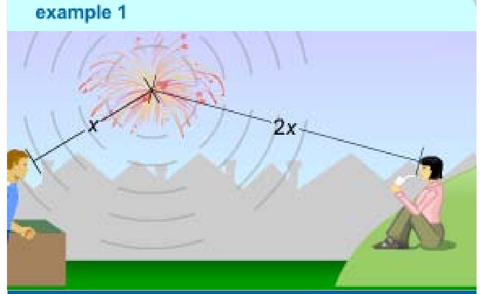


Sound spreading radially

$$I = \frac{P}{4\pi r^2}$$

P = power of sound source
r = distance from sound source

example 1



How many times more intense is the sound for the man than for the woman?

$$I = \frac{P}{4\pi r^2}$$

$$\frac{I_m}{I_w} = \frac{P/4\pi x^2}{P/4\pi(2x)^2}$$

$$\frac{I_m}{I_w} = \frac{(2x)^2}{x^2} = 4$$

17.9 - Sound level in decibels

Sound level: A scale for measuring the perceived intensity of sounds.

To compare the intensity of two sounds, you could directly calculate the ratio of their intensities. However, humans do not perceive a sound with twice the intensity as being twice as loud, so a different system may be used. Loudness is subjective, and having an objective measurement that corresponds to human perception is convenient. This measurement is called the *sound level* (or sometimes, more confusingly, the *intensity level*).

The human ear is sensitive to an extraordinary range of sound intensities; at the extremes, humans can perceive sounds whose intensities differ by a factor of 1,000,000,000,000. Although they can hear a broad range of intensities, people do not distinguish between them finely. For instance, the human ear cannot very well distinguish a sound that has an intensity of 1.0 W/m^2 from one with an intensity of 0.50 W/m^2.

To reflect humans' perception of differences in sound intensity, scientists use a logarithmic scale. The common unit for the sound level is the decibel (dB), or one-tenth of a "bel," a unit named after Alexander Graham Bell. A logarithmic scale provides an appropriate tool for describing the human perception of sounds.

In calculating a sound level β, you start by dividing the intensity of the sound being measured by a reference sound intensity that approximates the lowest intensity humans can hear. This reference intensity is 1×10^{-12} W/m^2. Then you calculate the common logarithm (to the base 10) of this ratio, which gives the sound level in bels, and finally you multiply that value by 10 to express the level in decibels. This is the first equation shown to the right.

To practice calculating sound levels in decibels, consider a sound with an intensity of 1.5×10^{-11} W/m^2. It has 15 times the sound intensity of the reference intensity of 1×10^{-12} W/m^2. The base-10 logarithm of 15 is 1.2. Multiplying this value by 10 decibels yields a sound level of 12 decibels. A sound level increase of 10 dB means the intensity increases by a factor of 10. In calculations of sound levels, both the numbers 1×10^{-12} W/m^2 and 10 (decibels) are considered exact.

You may note that the reference intensity of 1×10^{-12} W/m^2 corresponds to a decibel reading of zero. At zero decibels, a human ear can still barely hear sound. The pressure exerted by this sound is very, very slight: It displaces particles of air by about one hundred-billionth of a meter. It is possible to have negative decibel sounds, perturbations in air pressure so slight the human ear cannot detect them.

Tests indicate that a one to three decibel change in sound level is about the smallest change most humans can perceive. A general rule of thumb is that a human will perceive a tenfold increase in intensity as sounding twice as loud. A 50 dB sound is 10 times more intense than a 40 dB sound − remember, it is a logarithmic scale − and a typical human would say it sounds twice as loud.

The sound level equation can be recast in terms of sound power, as shown in the second equation to the right. If you know the relative power of two sound sources, you can use this equation to compare their relative loudness to the human ear, provided the listener is equidistant from the two sound sources. The reference power P_0 in this equation is the power at the source that results in the reference intensity at the location of interest. For instance, you might determine this value for a loudspeaker for an audience member 75 meters away.

concept 1

Sound	Sound level (decibels,dB)
Jet aircraft engine	160
Threshold of pain	140
Pneumatic drill	100
Subway train	90
Vacuum cleaner	85
Heavy auto traffic	75
Conversational speech	65
Whispered speech	40
Threshold of hearing	0

Sound level

Used to measure perceived loudness

· Units: decibels (dB)

· Logarithmic scale (20 dB is ten times more intense than 10 dB)

equation 1

Sound level

$$\beta = (10\text{ dB})\log\frac{I}{I_0}$$

β = sound level

I = intensity of sound

reference intensity $I_0 = 1\times10^{-12}$ W/m^2

$$\beta = (10\text{ dB})\log\frac{P}{P_0}$$

P = sound power of source

P_0 = reference sound power

Units: decibels (dB)

17.10 - Sample problem: sound level

Sitting on a sofa, your roommate hears 100 dB from your stereo, which supplies 10 W to each speaker. He says this is lame and requests a system with a maximum 120 dB. How many watts of power should the new stereo supply to each speaker?

Assume the sound power of the loudspeakers equals the power supplied by the stereo.

Variables

current power per speaker	$P_1 = 10$ W
reference power level	P_0
current sound level	$\beta_1 = 100$ dB
proposed power per speaker	P_2
proposed sound level	$\beta_2 = 120$ dB

What is the strategy?

1. State two equations that relate sound level to power, for both the current system and the proposed system.
2. Subtract the two equations and solve the resulting equation to determine the power of the new stereo.

Physics principles and equations

The equation for sound level with respect to the power of the sound source is

$$\beta = (10\text{ dB})\log\frac{P}{P_0}$$

Mathematics principles

$$\log(a) - \log(b) = \log(a/b)$$

Step-by-step solution

Step	Reason
1. $\beta_1 = (10\text{ dB})\log\frac{P_1}{P_0}$	sound level power equation
2. $100\text{ dB} = (10\text{ dB})\log\frac{10\text{ W}}{P_0}$	current system
3. $120\text{ dB} = (10\text{ dB})\log\frac{P_2}{P_0}$	proposed system
4. $20 = 10\log\frac{P_2}{P_0} - 10\log\frac{10}{P_0}$	subtract
5. $20 = 10\log\frac{P_2/P_0}{10/P_0}$	difference of logarithms
6. $2 = \log\frac{P_2}{10}$	divide and simplify
7. $10^2 = \frac{P_2}{10}$	take antilogarithm
8. $P_2 = 1000$ W	solve

To increase the sound level by 10 decibels, the sound power (and sound intensity) must increase tenfold. Raising the maximum sound intensity of the system by 20 decibels requires a 100-fold increase in sound power and stereo power. The result will be loud enough to ensure the entire dormitory hears your music!

We hope they appreciate your taste.

17.11 - Interactive checkpoint: coffee grinder sound level

A coffee grinder emits sound an all directions with power 5.05×10^{-4} W. At 0.400 meters from the grinder, what sound level do you hear?

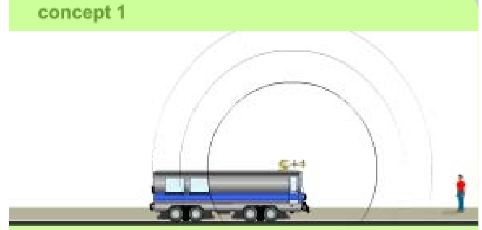
Answer:

β = ______ dB

17.12 - Doppler effect: moving sound source

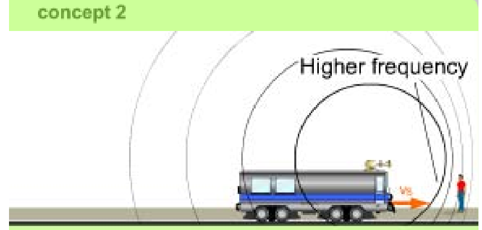
Doppler effect: A change in the frequency of a wave due to motion of the source and/or the listener.

concept 1



Sound source stationary
Wavelength, frequency constant

concept 2



Source moves toward listener
Sound wavefronts arrive closer together
Listener hears higher frequency

You experience the Doppler effect when a train races past you while sounding its whistle. As the train is approaching, you perceive the whistle as emitting sound of one frequency, and as it moves away, the perceived frequency of the whistle drops to a lower pitch.

This effect is named for the Austrian physicist who first analyzed it, Christian Doppler (1803-1853). Doppler's research concerned light from stars, but his principles apply to sound also.

In the example described above, the frequency of the sound emitted by the train is constant. The Doppler effect occurs because of the motion of the source of the sound, the train whistle. It is moving first toward and then away from you, and you are standing still. (What is moving and what is still is relative to sound's medium, the air.)

You see this situation illustrated on the right. In Concept 1, the train and listener are both stationary. The diagram shows the peaks of the sound waves as they emanate from the train and radiate in all directions, including toward the listener. They are equally spaced, which means their wavelength is constant, as is their frequency.

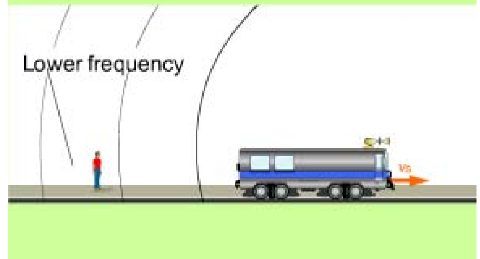
In Concept 2, the train is moving to the right, toward the listener, at a source velocity $\mathbf{v}_s$. As you see, the peaks of the sound waves at the listener are closer together than in Concept 1, which means the wavelength is shorter and the frequency at the listener is higher.

The motion of the train causes these changes in wavelength and frequency. To understand this, consider two successive regions of condensation generated by the train's horn. The first moves toward the listener. The train continues to move forward, and the next time the horn creates a region of condensation, it will be closer to the prior one than if the train were stationary. The regions arrive more frequently because of the motion of the train toward the listener.

Concept 3 shows the effect perceived by a listener for whom the train is moving away. The sound waves reach this listener less frequently, and he hears a lower pitched sound.

The Doppler effect is quantified using the two equations shown to the right. (They apply when the sound source is moving; a different set of equations is used when the listener is moving.) The first equation shows how to calculate the frequency when the source of the sound moves directly toward a stationary listener; the second is used when the source moves directly away. If the source is moving in some other direction, the component of its velocity directly toward or away from the listener must be used in the formulas. The speed of sound changes with temperature, air density and so on; a value of 343 m/s is often used.

concept 3



Source moves away from listener

Sound wavefronts arrive farther apart

Listener hears lower frequency

equation 1

Source moves toward listener

$$f_L = f_s \frac{1}{1 - v_s / v}$$

f_L = frequency perceived by listener

f_s = frequency emitted by source

v_s = speed of the source

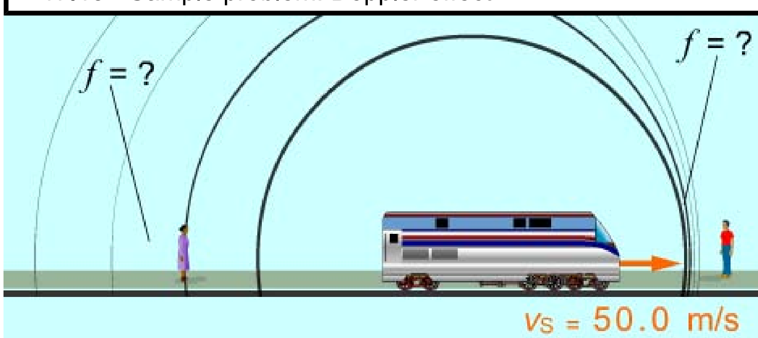
v = speed of sound in air ≈ 343 m/s

equation 2

Source moves away from listener

$$f_L = f_s \frac{1}{1 + v_s / v}$$

17.13 - Sample problem: Doppler effect



The frequency of the train whistle when the train is not moving is 495 Hz. What sound frequency does each person hear?

Variables

speed of train	v_s = 50.0 m/s
speed of sound	v = 343 m/s
frequency of train whistle when train is motionless	f_s= 495 Hz
frequency heard by front listener	f_{L1}
frequency heard by rear listener	f_{L2}

What is the strategy?

1. Use the equation for determining the Doppler effect when a sound-emitting object is moving toward a listener.
2. Then use the equation for determining the Doppler effect when a sound-emitting object is moving away from a listener.

Physics principles and equations

The equation for a sound source approaching a listener

$$f_L = f_s \frac{1}{1 - v_s / v}$$

The equation for a sound source moving away from a listener

$$f_L = f_s \frac{1}{1 + v_s / v}$$

Step-by-step solution

We start by calculating the frequency perceived by the front listener. The train is approaching that listener.

Step	Reason
1. $f_{L1} = f_s \frac{1}{1 - v_s / v}$	Doppler equation
2. $f_{L1} = 495\text{ Hz} \frac{1}{1 - \left((50.0\text{ m/s})/(343\text{ m/s})\right)}$	substitute values
3. $f_{L1} = 579\text{ Hz}$	evaluate

Now we calculate the frequency perceived by the rear listener, for whom the train is moving away.

Step	Reason
4. $f_{L2} = f_s \frac{1}{1 + v_s / v}$	Doppler equation
5. $f_{L2} = 495\text{ Hz} \frac{1}{1 + \left((50.0\text{ m/s})/(343\text{ m/s})\right)}$	substitute values
6. $f_{L2} = 432\text{ Hz}$	evaluate

This is quite a noticeable Doppler effect. The listener on the right hears a frequency roughly one-third higher than the listener on the left. If the train were playing a musical note, the listener on the right would hear a pitch about five semitones (piano keys) higher than the listener on the left.

17.14 - Doppler effect: moving listener or source

Three additional equations for calculating the Doppler effect are shown to the right. The first two are used to determine the effect when the sound source is stationary and the listener moves toward or away from it. The third − the grand finale of Doppler equations − calculates the effect when both the listener and the source are moving.

We start by using a non-sound example of mechanical waves to explain the effect when the listener is moving. Consider a boat anchored in a lake, with waves bumping into it at a constant rate. A boater sitting in the boat would perceive that waves are bumping into the boat at a constant frequency.

Now imagine that the boater pulls up the anchor, starts the engine and drives the boat into the oncoming waves. Because of the boat's motion, the frequency with which the boater meets the oncoming waves will increase. This is only the boater's perception. A swimmer resting at the edge of a dock would not observe any change in the frequency, wavelength or speed of the oncoming waves.

This effect is reversed if the boater drives directly away from the oncoming waves. In this case, the waves would bump into the boat less frequently, and the boater would perceive that the wave frequency had decreased.

Equation 1 applies when the listener moves toward a stationary sound source. The listener will perceive a higher frequency.

Equation 2 applies when the listener moves away from a stationary sound source, thereby perceiving a lower frequency.

Equation 3 covers all Doppler effect cases, as it incorporates variables for when both the sound source and the listener move relative to the air.

equation 1

Listener moves toward sound source

$$f_L = f_s\left(1 + \frac{v_L}{v}\right)$$

f_L = frequency perceived by listener
f_s = frequency emitted by source
v_L = speed of listener
v = speed of sound

Equation 3 contains plus/minus notation; it really is four equations in one. It covers the cases of the listener moving toward or away from the sound source and the sound source moving toward or away from the listener.

1. The sign is positive in the numerator when the **listener** moves in the direction **toward** the sound source. It is negative in the numerator when the **listener** moves **away** from the sound source.
2. The sign is positive in the denominator if the **sound source** moves in the direction **away** from the listener. It is negative in the denominator if the **sound source** moves **toward** the listener.

These two signs are independent, and all four combinations (+/+, +/−, −/+, −/−) are possible. Be careful: Remember that speed is always relative to the air. If the train is moving at −15 m/s along the track, and you are moving at −5 m/s as you chase it, you must use both these velocities in the equation in Equation 3. You cannot use a relative velocity of −10 m/s.

For all these equations, you can apply common sense as a reality check on whether you have the signs correct. Consider the physical situation and ask yourself this: Do you expect the frequency to increase or decrease for the listener?

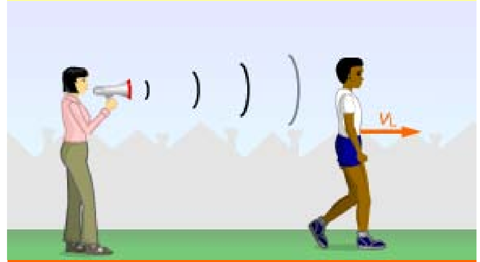
When the listener and sound source are moving apart, the perceived frequency for the listener decreases. The wave peaks encounter the listener less often. Make sure the signs you have chosen for the equation result in a decrease in the numerator or an increase in the denominator. That makes for a lower frequency.

When the listener and sound source approach each other, the frequency increases for the listener because the listener encounters wave peaks more often. A larger numerator, or smaller denominator, will make for an increased frequency.

These equations may seem complex, but certain types of bats use Doppler effects to locate and capture prey. Bats emit very high frequency (ultrasonic) sound waves in order to find their prey, flying insects. The waves bounce off the insects and return to the bat. Some bats, such as the horseshoe bat, analyze the Doppler shift of the returning sound waves to determine the relative motion of the insect: its speed and direction of travel. They also note how long the signal takes to return in order to calculate the distance to the insect. Using this information, the horseshoe bat flies from its perch to meet, and eat, the insect.

To counter the bats' sophisticated sonar system, some moths send out their own sound waves to try to jam the bats' signals. It is a competitive world out there!

equation 2



Listener moves from sound source

$$f_L = f_s\left(1 - \frac{v_L}{v}\right)$$

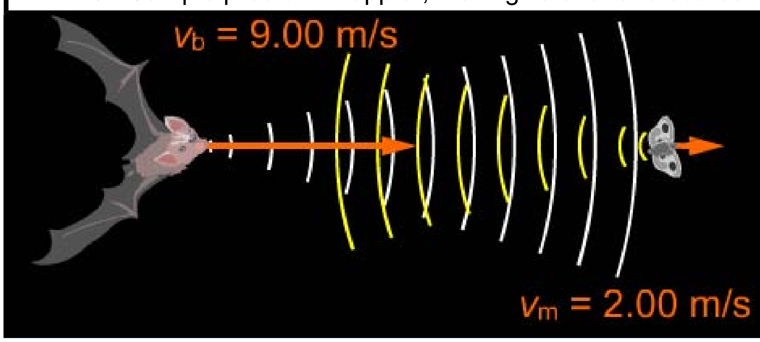
equation 3

Both listener and sound source move

$$f_L = f_s\left(\frac{1 \pm \frac{v_L}{v}}{1 \pm \frac{v_s}{v}}\right)$$

v_s = speed of source

17.15 - Sample problem: Doppler, moving listener and source



A bat flying 9.00 m/s emits a 65.0 kHz sound wave. It reflects off a moth fleeing at 2.00 m/s in the same direction the bat is flying. What frequency does the bat hear when the sound returns?

Variables

sound frequency emitted by bat	f_s = 65.0 kHz
frequency heard by moth	f_{Lm}
return frequency heard by bat	f_{Lb}
speed of bat	v_b = 9.00 m/s
speed of moth	v_m = 2.00 m/s
speed of sound	v = 343 m/s

What is the strategy?

1. This is a two-step problem. First, calculate the frequency received by the moth.
2. In the second phase of the problem, consider the effect when the sound reflects off the moth and returns to the bat. This makes the moth the sound source and the bat the listener.
3. In both steps, use the Doppler equation for when the listener and the sound source are both moving.

Physics principles and equations

The equation for when both listener and sound source are moving is

$$f_L = f_s\left(\frac{1 \pm \frac{v_L}{v}}{1 \pm \frac{v_s}{v}}\right)$$

Step-by-step solution

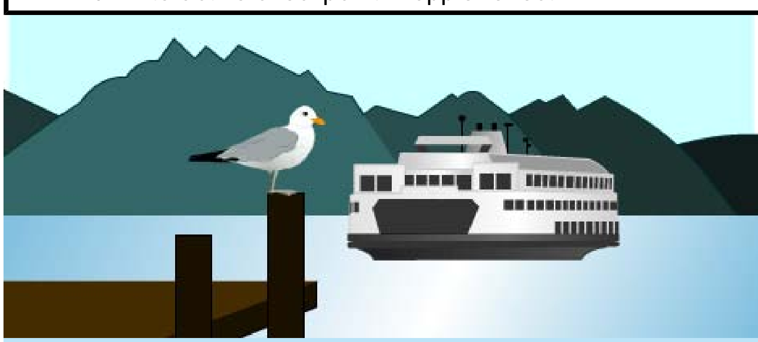
The first step is to determine the frequency of the sound waves reaching the moth. The bat and moth are both moving, so we need to use the equation that will factor in the motion of both source and listener. In this case, the bat is the source and the moth is the listener.

Step	Reason
1. $f_{Lm} = f_s\left(\frac{1 - v_m/v}{1 - v_b/v}\right)$	Doppler equation
2. $f_{Lm} = 65.0\text{ kHz}\left(\frac{1 - 2.00\frac{m}{s}/343\frac{m}{s}}{1 - 9.00\frac{m}{s}/343\frac{m}{s}}\right)$	substitute values
3. $f_{Lm} = 66.4\text{ kHz}$	solve

Now the sound wave reflects off the moth and travels back to the bat. In this step, the moth is the sound source and the bat is the listener. We use the frequency value calculated above as the frequency of the source sound.

Step	Reason
4. $f_{Lb} = f_{Lm}\left(\frac{1 + v_b/v}{1 + v_m/v}\right)$	Doppler equation
5. $f_{Lb} = 66.4\text{ kHz}\left(\frac{1 + 9.00\frac{m}{s}/343\frac{m}{s}}{1 + 2.00\frac{m}{s}/343\frac{m}{s}}\right)$	substitute values
6. $f_{Lb} = 67.7\text{ kHz}$	solve

17.16 - Interactive checkpoint: Doppler effect



A ferry pulls directly away from a dock while emitting sound of frequency 198 Hz. A bird sitting on the dock hears a frequency of 191 Hz. At what speed is the ferry moving? Use 343 m/s for the speed of sound.

Answer:

v_s = ______ m/s

17.17 - Interactive problem: a bat on the hunt

A horseshoe bat is perched upside down on a tree limb, as shown in the diagram to the right. A particularly clueless moth is flying directly at the bat. If the bat can determine when to open its mouth, it will not have to move an inch to enjoy a yummy snack.

Help the bat get its lunch. In this interactive problem, you reflect bat sonar off the moth to determine (a) the moth's location and (b) the moth's flying speed. Taken together, these two facts will enable you to time when to open and close the bat's mouth to allow it to swallow the moth.

In the simulation, press GO to start the moth flying across the screen toward the bat. At any time, press the SEND SONAR button. High frequency sound waves will travel out from the bat, reflect off the moth, and return to the bat.

The simulation will then pause, and you will see displayed in the control panel the total time, in milliseconds, it took the sound wave to travel to the moth and back. (The simulation runs in slow motion, so this is the not elapsed time you will experience.) Use this data to calculate the distance from bat to moth. The speed of sound in the simulation is 336 m/s (it is a little chilly out tonight).

Also displayed in the control panel is the frequency of the sound waves emitted by the bat, as well as the frequency of the sound waves that return to the bat. Use this data and the Doppler effect equations to determine the speed of the flying moth.

Using the Doppler equation is not a trivial undertaking for you or the bat. Note that because the moth is moving, the sound is Doppler shifted twice: once when the sound wave traveling from the bat reaches the moving moth, and again when the reflected sound wave returns to the bat. In the first instance, consider the bat as a stationary sound source and the moth as a moving listener. In the second instance, consider the moth as a moving sound source and the bat a stationary listener. Because the sound is Doppler-shifted twice, you must apply two separate Doppler equations to find the speed of the flying moth.

Once you have determined the moth's distance from the bat and the speed at which it is flying, decide when the bat should open its mouth to catch the moth. Enter the time, in seconds, in the box in the control panel. Press GO to resume the simulation, and the moth will continue to fly toward the bat.

You have succeeded if the bat swallows the moth. Otherwise, press RESET and try again.

If you have trouble determining the moth's distance from the bat, review the sections on the speed of sound in air. Note that the moth moves slightly closer to the bat in the time it takes the sound wave to return to the bat after reflecting off the moth. Since the moth moves much more slowly than the speed of sound, the distance is so slight that you can ignore it in your calculations.

If you have trouble determining the moth's flying speed, review the sections on the Doppler effect. Pay particular attention to the sample problem on Doppler shifts in a reflected sound wave.

17.18 - Supersonic speed and shock waves

After studying the Doppler effect equations, you might wonder: What happens if the speed of the sound source equals the speed of sound? The relevant Doppler equation shows that you would have to divide by zero, yielding infinite frequency, a troubling result.

Equally troubling was the result when aircraft first tried to break the *sound barrier* (the speed of sound). In the first half of the 20th century, the planes tended to fall apart as much as the equation does. It was not until 1947 that a plane was able to fly faster than the speed of sound, as excitingly shown in the movie *The Right Stuff*, proving that sound was not a barrier after all.

This boat is traveling faster than the speed of waves in water. Its wake forms a two-dimensional "Mach cone" on the water's surface.

The Doppler effect equations do not apply if the sound source is moving at or above the speed of sound. Concept 1 shows the result when a sound source moves at the speed of sound. The leading edges of the sound waves bunch up at the tip of the aircraft as the plane travels as fast as its own sound waves.

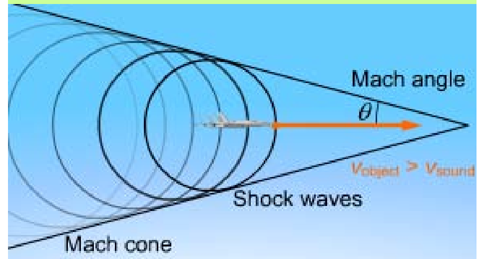
Concept 2 shows the result when an aircraft exceeds the speed of sound. Aircraft capable of flying that fast are called *supersonic*. The plane travels faster than its own sound waves, and the waves spool out behind the plane creating a *Mach cone*.

The surface of the Mach cone is called a *shock wave*. Supersonic jets produce shock waves, which in turn create sounds called *sonic booms*. As long as a jet exceeds the speed of sound, it will create this sound. A rapid change in air pressure causes the sonic boom. Shock waves may be visible to the human eye because a rapid pressure decrease lowers temperature and causes water molecules to condense, resulting in fog. You may have heard other sonic booms, such as the report of a rifle or the crack of a well-snapped whip. The boom indicates that the bullet or the tip of the whip has moved faster than the speed of sound. The example of the whip shows that the moving object can be silent and still create a shock wave.

When object is at speed of sound
It travels as fast as its own sound waves

Equation 1 shows how the sine of the angle of a Mach cone can be calculated as a ratio of speeds. The inverse ratio, of the speed of the object to the speed of sound, is known as the *Mach number*. A fighter jet described as a "Mach 1.6 plane" can move as fast as 1.6 times the speed of sound, or about 550 m/s (more than 1200 mph). You see this described mathematically in Equation 2. Because the speed of sound varies with factors like temperature, the exact speed of a Mach 1.6 plane depends on its environment.

concept 2



Supersonic speed

Object exceeds speed of its sound waves

· Sound waves form Mach cone
· Surface of cone is shock wave
· Angle θ is called Mach angle

equation 1

Mach angle

$$\sin\theta = \frac{v_{\text{sound}}}{v_{\text{object}}}$$

θ = Mach angle
v_{sound} = speed of sound
v_{object} = speed of object

equation 2

Mach number

$$\text{Mach number} = \frac{v_{\text{object}}}{v_{\text{sound}}}$$

17.19 - Gotchas

You hear sound because air molecules move from the sound's source to your ears. No. Air molecules oscillate back and forth as a sound wave passes by, but there need be no net motion of the particles that make up the medium.

The speed of sound can vary significantly depending on the medium it travels through. Yes. The speed of sound varies widely. Sound travels nearly five times faster through water than through air, for instance, and faster still through solids, including many metals.

The greater the amplitude of a sound wave, the faster the wave moves. No. The speed of a sound wave is dependent solely on the properties of the medium through which it moves.

A 20 dB sound is twice as intense as a 10 dB sound. No, the decibel measurement system is logarithmic. The 20 dB sound has ten times the intensity: ten times as many watts per square meter. A typical person perceives it as twice as loud.

The degree of Doppler shift experienced by a listener is independent of how far the listener is from a moving sound source. This is true: Distance is not a factor in determining the Doppler shift.

17.20 - Summary

Sound waves are longitudinal mechanical waves. Instead of the peaks and troughs of transverse waves, sound waves are composed of condensations and rarefactions of the medium through which they travel. Particles in a sound wave move in simple harmonic motion along the direction that the wave travels.

The speed of sound depends upon the properties of the medium through which it travels. In air, the speed increases as temperature increases. In fluids and solids, the speed of sound is proportional to the modulus of elasticity and inversely proportional to the density.

One way to mathematically describe a sound wave is to consider the sinusoidal

Equations

Speed of sound in air

$$v = (331\text{ m/s})\sqrt{1 + \frac{T_C}{273\ {}^\circ\text{C}}}$$

Speed of sound in fluids

motion of a particle as it transmits the wave. Another way is to focus on how the air pressure changes as the wave passes by. The change in air pressure is also a sinusoidal function. The amplitude of the pressure-difference function is a function of the wave's amplitude and frequency, the density of the air, and the wave's speed.

The intensity of a sound is the sound power passing perpendicularly through a unit area. For a sound that spreads radially, such as from a fireworks explosion in midair, the sound intensity is inversely proportional to the square of the distance from the source.

Do not confuse sound intensity with the sound level. The sound level takes into account the logarithmic perception of sound by the human ear.

The Doppler effect is the change in frequency of a wave due to the relative motion of the source and/or the listener. A common example is the change in frequency heard as the siren on a fire engine races by you. As the fire engine moves toward you, the sound waves "pile up" and their wavelength decreases. Since the speed of sound remains the same, this results in you hearing a higher frequency.

$$v = \sqrt{\frac{B}{\rho}}$$

Speed of sound in solids

$$v = \sqrt{\frac{Y}{\rho}}$$

Sound wave displacement

$$s(x,t) = A\cos\left(\frac{2\pi x}{\lambda} - 2\pi ft\right)$$

Sound wave pressure

$$\Delta P(x,t) = \Delta P_{max}\sin\left(\frac{2\pi x}{\lambda} - 2\pi ft\right)$$

$$\Delta P_{max} = 2\pi Af\rho v$$

Sound intensity

$$I = P/A$$

Sound spreading radially

$$I = \frac{P}{4\pi r^2}$$

Sound level

$$\beta = (10 \text{ dB})\log\frac{I}{I_0}$$

Doppler effect, source moves toward listener

$$f_L = f_s\frac{1}{1 - v_s/v}$$

Doppler effect, source moves away from listener

$$f_L = f_s\frac{1}{1 + v_s/v}$$

Doppler effect, listener moves toward source

$$f_L = f_s\left(1 + \frac{v_L}{v}\right)$$

Doppler effect, listener moves away from source

$$f_L = f_s\left(1 - \frac{v_L}{v}\right)$$

General Doppler effect

$$f_L = f_s\left(\frac{1 \pm \frac{v_L}{v}}{1 \pm \frac{v_s}{v}}\right)$$

Mach angle

$$\sin\theta = \frac{v_{\text{sound}}}{v_{\text{object}}}$$

$$\text{Mach number} = \frac{v_{\text{object}}}{v_{\text{sound}}}$$

Chapter 17 Problems

Chapter Assumptions

Unless stated otherwise, use 343 m/s for the speed of sound.

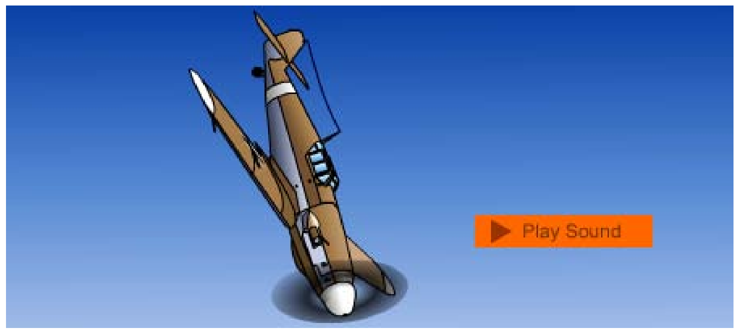
Conceptual Problems

C.1 An old trick to find out if a train is coming is to put your ear on the rails of the track. Why is this more effective than just listening for the sounds the train makes in air?

C.2 Why does the sound of a jet plane flying overhead seem to come from a direction trailing the plane itself?

C.3 Why do people at public events use cone-shaped bullhorns to address the crowd? (We are referring to the non-battery-powered type!)

C.4 If you have ever watched a World War II movie, you may have heard the frightening sound made by a German Stuka dive bomber as it dove straight down toward its target on the ground during a bombing run. This falling sound, which may be approximated in writing as "eeeeaaaaooooowwww", was produced by a single-note siren attached to each bomber as an instrument of psychological terror by the Luftwaffe. However, a GI who was knowledgeable about physics would actually have been much more frightened to hear an unvarying high-pitched "eeeeeeeeeeeeeeeee" during the plane's vertical dive. Why is this? (Note: acceleration due to gravity will not be a factor in your explanation. The plane is diving at its propeller-driven terminal velocity.)



C.5 It is a lovely day, sunny, with a nice steady breeze blowing. As you stand, dreaming, you hear the tinkle of wind chimes carried to you from directly upwind. Is the frequency of the tinkles you hear higher, the same as, or lower than the frequency emitted by the chimes?

i. Higher
ii. The same
iii. Lower

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Does the pitch get higher, lower or does it stay the same as the frequency increases? (b) Does the pitch get higher, lower or does it stay the same as the amplitude increases?

(a) i. Higher
ii. Lower
iii. Stays the same

(b) i. Higher
ii. Lower
iii. Stays the same

Section 3 - Interactive problem: sound frequency

3.1 Use the simulation in the interactive in this section to determine whether higher pitched notes have higher or lower frequencies than lower pitched notes.

i. Higher
ii. Lower
iii. The same

Section 4 - Speed of sound in various media

4.1 Any longitudinal wave in water (such as the shock wave from an explosion) moves at the same speed as sound in water. Calculate the speed of longitudinal waves in water. The bulk modulus *B* of water is 2.18×10^9 N/m^2 and its density is 1.00 g/cm^3 at 4°C.

__________ m/s

4.2 The highest natural atmospheric temperature ever recorded on Earth was 58°C (136°F), at El Azizia, Libya on September 13, 1922. The record low temperature was −89°C (−129°F), which occurred at the Soviet Vostok Station in Antarctica on July 21, 1983. What is the difference in the speed of sound in air for these two extreme temperatures?

__________ m/s

4.3 You are viewing a New Year's Day fireworks display at a distance of 1.3e+3 m. It is 4.0°C. How long does it take the sound of the fireworks exploding to reach you?

__________ s

4.4 Calculate the speed of sound in cast steel, which has a Young's modulus of 197×10^9 N/m^2 and a density of 7750 kg/m^3.

__________ m/s

4.5 Seawater has a bulk modulus of 2.34×10^9 N/m^2. Its density at 5°C is 1030 kg/m^3. What is the speed of sound through seawater at this temperature?

__________ m/s

4.6 Sound travels at 1066 m/s through Ethyl alcohol (ethanol) at room temperature. Its density at this temperature is 789 kg/m^3. What is its bulk modulus?

__________ N/m^2

4.7 Young's modulus for copper is 108×10^9 N/m^2. Sound travels at 3480 m/s in copper at room temperature. What is its density at the same temperature?

__________ kg/m^3

Section 7 - Mathematical description of a sound wave

7.1 The threshold of pain for the human ear corresponds to an overpressure amplitude of 29.0 Pa. What is the displacement amplitude for air at this overpressure amplitude? Use a frequency of 1000 Hz, an air density of 1.21 kg/m^3 and the speed of sound in air at room temperature.

__________ m

Section 8 - Sound intensity

8.1 You are standing 7.0 meters from a sound source that radiates equally in all directions, but it is too loud for you. How far away from the source should you stand to experience one third the intensity that you did at 7.0 meters?

__________ m

8.2 15.0 meters from a sound source that radiates freely in all directions, the intensity is 0.00460 W/m^2. What is the rate at which the source is emitting sound energy?

__________ W

8.3 A longitudinal wave spreads radially from a source with power 345 W. What is the intensity 40.0 meters away?

__________ W/m^2

8.4 Sound spreads radially in all directions from a source with power 15.3 W. If the intensity you experience is 3.00×10^{-6} W/m^2, how far away are you from the source?

__________ m

8.5 As the featured speaker at a rally, you yell through an unpowered bullhorn (basically, a cone) to excite the crowd. The cone directs your voice, and is defined by a half-angle of 0.45 radians. Since sound expands radially, the wave fronts are sections of a sphere. The surface area of the section of a sphere of radius r defined by a cone with a half-angle θ is given by $2\pi r^2(1 - \cos\theta)$. (a) Calculate the ratio of sound intensity experienced by people 2.0 meters from you to that experienced by people located 4.0 meters from you. (b) For people at a given distance *r,* what is the ratio of the sound intensity they hear to the sound intensity they would hear if you were broadcasting at the same power with no cone, so that the sound waves spread spherically and symmetrically in all directions?

(a) __________
(b) __________

Section 9 - Sound level in decibels

9.1 The sound of heavy automobile traffic, heard at a certain distance, has a sound level of about 75 dB. What is the intensity of the noise there?

__________ W/m^2

9.2 Jane and Sam alternately pound a railroad spike into a tie with their hammers. The crew chief has a migraine, and notes that Jane's hammer blows cause a sound with intensity 5.0 times greater than the sound that Sam makes when he swings his hammer. What is the difference in sound level between the two sounds?

__________ dB

9.3 A grade-school teacher insists that the the overall sound level in his classroom not exceed 64 dB at his location. There are 25 students in his class. If each student talks at the same intensity level, and they all talk at once, what is the highest sound level at which each one can talk without exceeding the limit?

__________ dB

9.4 A cafe is known for its tasty blended juice drinks. The owner worries that if too many blenders are running in the preparation area at once, the noise will drive customers away. She determines that the maximum allowable sound level at the cash register should be 84 dB. If each blender in the cafe emits sound with a level of 79 dB at the register, how many blenders can be running at once before the maximum is exceeded?

__________ blenders

9.5 The human eardrum has an area of $5.0\times10^{-5}\ \mathrm{m}^2$. (a) What is the sound power received when the incident intensity is the minimum perceivable, $1.0\times10^{-12}\ \mathrm{W/m}^2$? (b) What is the sound power at the eardrum for conversational speech, which has a sound level of 65 dB? (c) What is the sound power at the threshold of pain, 140 dB?

(a) __________ W
(b) __________ W
(c) __________ W

9.6 Pete's family is watching a movie at home, with the TV emitting sound at a level of 71 dB as heard by the family. His dad turns on the popcorn popper, which emits sound that has a level of 68 dB as heard by the family. (a) What is the total sound intensity perceived by Pete's family? (b) What is the total sound level?

(a) __________ W/m^2
(b) __________ dB

9.7 Suzy revs her monster truck, producing a sound with a sound level of 84 dB as heard by Mrs. DiNapoli across the street. Next door to Suzy, Mike guns his station wagon which produces a sound with a sound level of 71 dB at Mrs. DiNapoli's. At this poor neighbor's house across the street, what is the unbearable ratio of the intensity of the truck's sound to the intensity of the wagon's sound?

____

9.8 Ju Yeon sits in a park where people come to fly small radio-controlled airplanes. She can hear down to a sound level of 20 dB. If a plane emits sound with a power of 1.6e−6 W, how far away will the plane be when she is just able to hear it?

__________ m

Section 12 - Doppler effect: moving sound source

12.1 Passengers on a train hear its whistle at a frequency of 735 Hz. What frequency does someone standing by the train tracks hear as it moves directly toward her at a speed of 22.5 m/s?

__________ Hz

12.2 A particular ambulance siren alternates between two frequencies. The higher frequency is 617 Hz as heard by the ambulance driver. If the ambulance drives directly away from you at a speed of 16.8 m/s, what frequency do you hear as the higher frequency?

__________ Hz

12.3 A small plane is taxiing directly away from you down a runway. The noise of the engine, as the pilot hears it, has a frequency 1.13 times the frequency that you hear. What is the speed of the plane?

__________ m/s

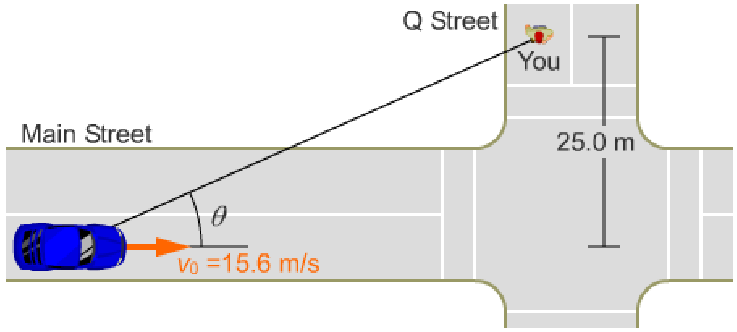
12.4 A train moves away from you at 29.0 m/s, sounding its horn. You perceive the frequency of the horn's sound as being 415 Hz. What is the frequency heard by the train's passengers?

__________ Hz

12.5 A stuntwoman is preparing to take a punch, crash through a "candy glass" window, and fall a long distance. The script calls for her to emit a piercing scream just before she hits the "ground." In reality, she will land on a waiting airbag. Lights! Camera! Action! The primary camera crew, filming from her starting height, hears her last-instant scream at a frequency 3.77 kHz. Her scream has a frequency of 4.05 kHz when she is at rest. How far did she fall? Report this as a positive distance.

__________ m

12.6 Q Street (which runs north-south) is perpendicular to Main Street (which runs east-west). You are standing on Q Street, 25.0 meters north of its intersection with Main. A car drives east on Main at v_0 = 15.6 m/s, sounding its horn at 435 Hz, as heard by the driver. (a) Does the frequency you hear increase or decrease as the car approaches the intersection? (b) The car sounds its horn when it is 12.4 meters from the intersection. What frequency do you hear? (c) The car sounds its horn when it is 2.53 meters from the intersection. What frequency do you hear?



(a) ○ Increase ○ Decrease
(b) __________ Hz
(c) __________ Hz

Section 14 - Doppler effect: moving listener or source

14.1 A team of tornado-chasing meteorology graduate students is out in search of tornados. They drive directly toward a tornado warning siren at 30.0 m/s. The siren emits sound at 685 Hz. What frequency of sound do the students hear?

__________ Hz

14.2 At a family picnic, John's annoying nephew sits under a tree blowing into a penny whistle. John runs away from the obnoxious noise at 3.59 m/s and hears a sound of frequency 845 Hz. What frequency does his nephew hear?

__________ Hz

14.3 Takashi is riding a Japanese bullet train, or Shinkansen, that is heading away from Tokyo at 0.250 times the speed of sound. A Tokyo-bound Tohoku Shinkansen is coming toward it on an adjacent track, carrying Yuki at 0.113 times the speed of sound. Acting upon a very unwise impulse, Takashi and Yuki stick their heads out their respective windows and emit primal screams at 3.50 kHz. (a) At what frequency does Takashi hear Yuki's scream? (b) At what frequency does Yuki hear Takashi's scream?

(a) __________ kHz
(b) __________ kHz

14.4 You are driving down a road at a constant speed of 17.8 m/s and another car drives toward you. (The other car is in the other lane, but you can ignore the slight lateral difference in the paths and treat the problem as if the car drives directly toward you. Or else assume that you are foolishly playing "chicken" and it **is** driving directly toward you.) The other car has a constant speed of 20.1 m/s from the viewpoint of someone standing by the side of the road. If you sound your horn at 455 Hz, what frequency do the people in the other car hear?

__________ Hz

14.5 Car A and car B are driving in the same direction. Car A's speed is 1.50 times that of car B. Car A sounds its horn at a frequency of 466 Hz and the driver of car B hears a frequency of 476 Hz. (a) Which car is ahead, A or B? (b) How fast is car B moving?

(a) ○ A ○ B
(b) __________ m/s

Section 15 - Sample problem: Doppler, moving listener and source

15.1 A dolphin swims toward a boat at 9.60 m/s. The boat heads toward the dolphin at a speed of 6.70 m/s. If the dolphin emits a call at 120 kHz, what frequency does it hear after the sound reflects off the boat? The speed of sound in seawater is 1522 m/s.

__________ kHz

Section 17 - Interactive problem: a bat on the hunt

17.1 Use the information given in the interactive problem in this section to determine the speed of the moth.

________ m/s

Section 18 - Supersonic speed and shock waves

18.1 The Mach angle of a jet traveling at supersonic speed is 19.4°. What is the Mach number?

____

18.2 What is the Mach angle of a jet moving at Mach 2.2?

________ °

18.3 A jet travels at a speed of 470 m/s through air at 0°C. What is the Mach number?

____

18.4 A jet fighter flies directly overhead. After it has swept out an additional 65.0° through the sky (that is, it is 25.0° above the horizon) the sonic boom hits you. Assuming that the plane is low enough that the air temperature does not change drastically, how fast is the plane moving? Use an air temperature of 6.00°C.

________ m/s

18.5 A jet plane flies directly overhead at Mach 1.50 at an altitude of 9120 meters. Assume that the air between you and the plane is a constant 0.00°C. (a) How far horizontally does the plane travel before its shock wave reaches you? (b) How long after the plane is directly overhead do you hear it?

(a) ________ m
(b) ________ s

Additional Problems

A.1 While kayaking in a lake you paddle directly away from the shore at a speed of 0.41 m/s. Waves that are traveling toward the shore at 0.96 m/s (according to the perspective of a shore observer) hit your kayak with a frequency of 2.0 waves per second. What is the frequency of the waves hitting the shore?

________ Hz

A.2 You experience the sound of a motor running at a sound level of 66 dB when you are 11 meters away. What is the sound power emitted by the motor? Assume it radiates sound equally in all directions.

________ W

A.3 You note that as a motorcycle approaches you directly and then passes you, the noise that its engine makes decreases in frequency, so that it is 0.84 times the initial frequency. What is the speed of the motorcycle? Neglect the small lateral distance between you and the motorcycle.

________ m/s

A.4 A boat sounds a foghorn which is heard by a lighthouse operator at a sound level of 54 dB. (a) The lighthouse operator knows that the sound power of the boat's horn is 4.5 W. How far away is the boat? Assume that the horn is a sound source that radiates equally in all directions, and that the fog does not absorb any sound. (b) For someone standing on the boat, 7.5 meters from the horn, what is the sound level?

(a) ________ km
(b) ________ dB

A.5 For the same sound source that radiates equally in all directions (an *isotropic* sound source), the difference in sound levels received at two distances r_1 and r_2 is given by $\beta_2 - \beta_1 = x \log (r_1/r_2)$. In this formula, the multiplier x is a constant that does not depend on the radii. Find the value of x in decibels.

________ dB

A.6 Two trains are moving directly away from one another. One is bound for Seattle, Washington and moves at a constant speed of 16.0 m/s. The other is bound for Vancouver, Washington and moves at 10.1 m/s. The train bound for Vancouver sounds its horn at a frequency of 423 Hz and the Seattle-bound train hears a frequency of 392 Hz. (a) What is the speed of sound in the area that day? (b) What is the air temperature?

(a) ________ m/s
(b) ________ ° C

A.7 A train starts from rest at your stationary position and heads directly away from you. It accelerates at a constant 0.50 m/s^2. The conductor blows the horn, which she perceives as having a frequency of 455 Hz. How far away is the train at the moment when you hear a frequency of 435 Hz? Hint: Remember that it takes time for the sound of the train's horn to reach you.

__________ m

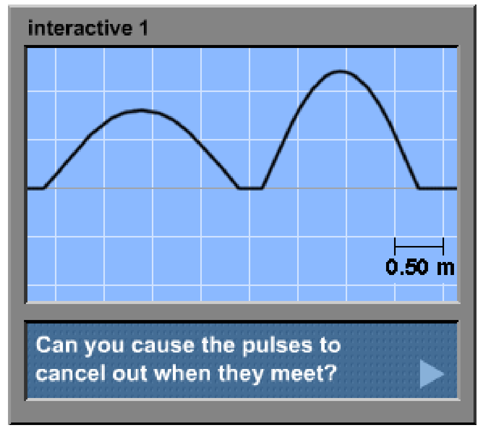
Chapter 18 Wave Superposition and Interference

18.0 - Introduction

In this chapter, we will discuss what happens when two or more traveling waves combine with each other, as when waves meet in a pond, a pool or even a bathtub. The result is higher peaks and lower troughs. These waves can “pass through” each other and then regain their original shape and direction, in contrast to the collision of two particles, such as tennis balls, which alters the motion of the balls.

Intersecting ripples from two different wave sources in a pond.

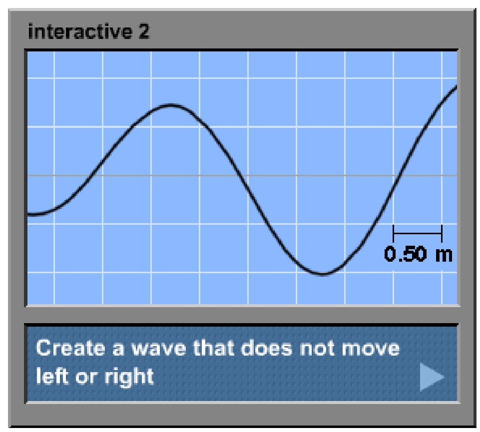
A stringed musical instrument like a violin or a guitar uses the reflection and recombination of waves from the end of a bowed or plucked string to create its magical sound. You also hear the result of waves combining when you listen to music on loudspeakers or as a live performance of a band or a symphony. Music halls are designed to take advantage of the reflection and combination of sound waves to produce an effect that audience members will find particularly pleasing.



You can begin your study of the result of combining waves with the simulations to the right. In Interactive 1, you control two wave pulses traveling on the same string. One pulse starts on the left and moves right, and the other starts on the right and moves left. You determine the amplitude and width of each pulse, as well as whether it is a peak or a trough. Set the parameters for each pulse and press GO to see what happens when the pulses meet. If you want to see the pulses combining in slow motion, press the “time step” arrow to advance or reverse time in small increments.

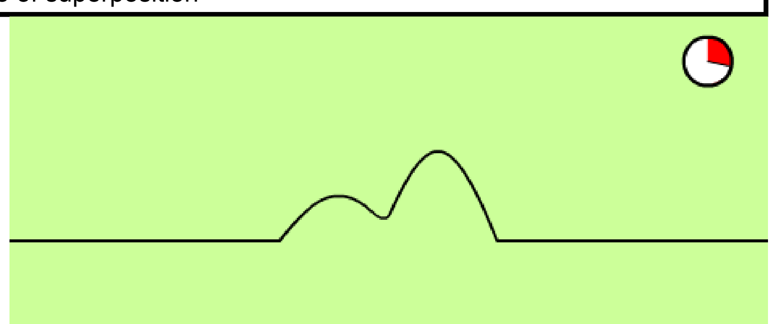
A challenge for you: Can you set the pulses so that when they meet they exactly cancel each other out, causing the string to be completely flat for an instant?

In Interactive 2, two traveling waves combine as they move toward one another. You determine the amplitudes and wavelengths of these waves. Again, set the parameters of the waves and press GO to see the result of combining the waves. Can you create a single combined wave that does not move either to the left or to the right? If you do, you will have created what is called a standing wave. You will find that by making the settings of the two waves identical, except for travel direction, you can create such a wave.



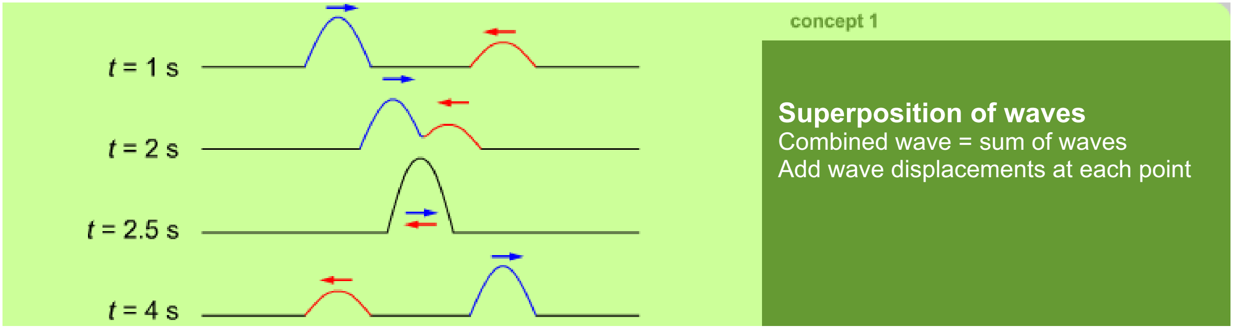
18.1 - Combining waves: the principle of superposition

In much of this chapter, we discuss what happens when traveling waves combine. In this section, we consider the less complicated case of what occurs when two wave pulses combine in a string, as you see above. This will help us illustrate the principle of superposition, which is more readily viewed with pulses than with traveling waves.



Two traveling wave pulses on a string.

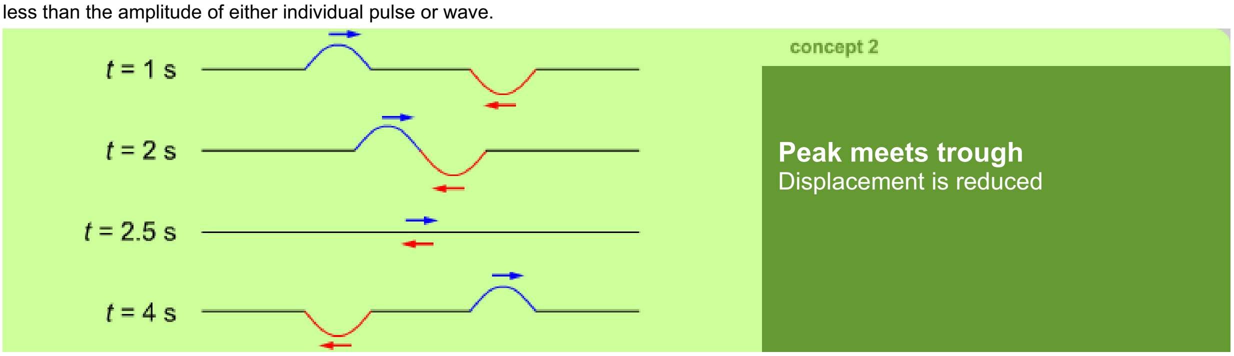
In Concept 1 below, we show what occurs when two peaks like the ones in the illustration above combine. We show the result at four successive times. A blue pulse is traveling from left to right and a red pulse is traveling in the opposite direction. You can see that at each instant the combined pulse is determined by adding the vertical displacements of the two pulses at every point along the string.



In the description above, we are applying the *principle of superposition*: The wave that results when two or more waves combine can be determined by adding the displacements of the individual waves at every point in the medium. This is sometimes more tersely stated as: The resulting wave is the algebraic sum of the displacements of the waves that cause it. ("Algebraic" means that you add positive and negative displacements as you would any signed numbers; no trigonometry is required.)

You see this principle in play in Concept 1. For instance, when the two peaks meet at time $t = 2.5$ s, the resulting peak's displacement equals the sum of the displacements of the two separate peaks. This is an example of *constructive interference*, which occurs when the amplitude of a combined pulse or wave is greater than the amplitude of any individual pulse or wave.

In Concept 2 below, a peak meets a trough. Except for the different directions of displacement, the pulses are identical. The two pulses cancel out completely when they occupy the same location on the string, and the string momentarily has zero displacement at each point. Positive and negative displacements of the same magnitude sum to zero. This is *destructive interference*: The amplitude of the combined pulse or wave is less than the amplitude of either individual pulse or wave.

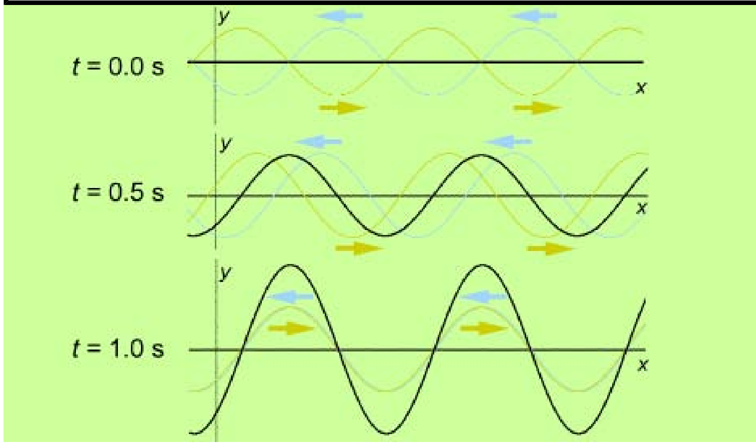


You may have experimented with this in the introduction simulation, but if you did not, you can go back and see what happens when peak meets peak, when peak meets trough, and finally, when two troughs meet. The result in each case is that the combined displacement is the sum of the displacements of all the pulses.

When the string is "flat" in Concept 2, it may seem there is no motion because you are looking at a static diagram. In fact, some of the string particles are moving up and some are moving down. The particles that were part of a peak are moving down and will be part of a trough. This is readily witnessed in the introductory simulation.

In this section, we used transverse wave pulses on a string to illustrate superposition, but the principle of superposition can also be applied to longitudinal waves (for instance, the combined sound wave produced by two stereo speakers). Acoustical (sound) engineers rely on constructive interference to create louder sounds and destructive interference to mask noises. For instance, noise-reducing headphones contain a microphone that detects unwanted noise from the environment. A circuit then creates a sound wave that is an inverted version of the noise wave, with peaks where the noise has troughs, and vice-versa. When this wave is played through the headphones, it destructively interferes with the unwanted ambient noise. The same technique is used to reduce the noise from fans in commercial heating and ventilation systems.

18.2 - Standing waves



concept 1

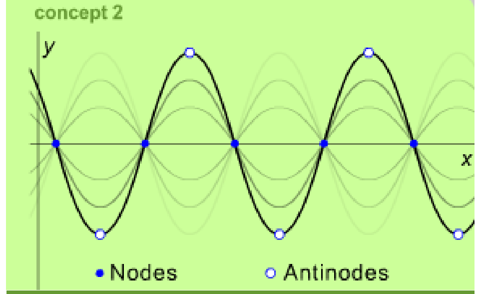
Standing wave
Created by identical waves moving in opposite directions

Standing wave: Oscillations with a stationary outline produced by the combination of two identical waves moving in opposite directions.

When two identical traveling waves move in opposite directions through a medium, the resulting combined wave stays in the same location. The resulting wave is called a standing wave. Individual particles oscillate up and down, but unlike the case of a traveling wave, the locations of the peaks and troughs of a standing wave stay at fixed positions.

Consider the illustration above, showing waves on a string. The standing wave is formed by the combination of two traveling waves. We show three "snapshots" in time of two identical waves traveling in opposite directions, and the combined wave they create. The traveling waves are shown in colors. These waves have the same frequency and amplitude, but are traveling in opposite directions. The result is a standing wave that does not move along the string.

concept 2



Standing wave
Nodes: no displacement
Antinodes: maximum displacement

In the first snapshot above, the two traveling waves are out of phase, and they destructively interfere. The combined standing wave has a constant zero displacement. In the second snapshot, the traveling waves have moved slightly, and the standing wave exhibits some peaks and troughs. In the third snapshot, the traveling waves exactly coincide, in phase, to constructively interfere, and the standing wave has its largest peaks and troughs. As the traveling waves continue to move, the standing wave's peaks and troughs will diminish until the traveling waves are again out of phase and the standing wave is flat.

If you find this difficult to visualize, you can see a simulation of the creation of a standing wave by clicking on the whiteboard above. You can have the simulation move slowly, step by step, by pressing the arrow buttons. You can also press "show components" if you want to see the component traveling waves. Press REPLAY to restart the simulation.

Along a standing wave, there are some fixed positions where there is no displacement and others where there is the maximum displacement. The positions with no displacement are called *nodes*. The positions where the wave has the greatest displacement (the peaks and troughs) are called *antinodes*. Adjacent nodes (and antinodes) are separated by a constant distance. Looking at the illustration for Concept 2, you can see that two adjacent nodes are separated by one half the wavelength of the wave. The same is true for two adjacent antinodes.

The standing wave depicted above is created by the combination of transverse waves. Longitudinal waves, such as sound waves, can combine to form standing waves as well.

18.3 - Standing wave equations

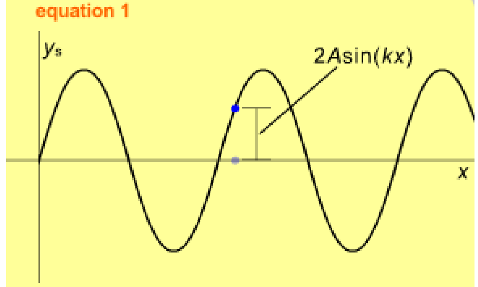
A standing wave results when two waves with the same amplitude and wavelength travel in opposite directions in the same medium. The first two equations in Equation 1 describe two such traveling waves. The first equation describes a wave traveling to the left, because ωt is positive, while the second describes a wave that is traveling to the right, because ωt is negative in that equation.

We can apply the principle of superposition and add the two functions to determine the combined wave. When we add them, we use the trigonometric identity

$$\sin a + \sin b = 2\sin\left(\frac{a+b}{2}\right)\cos\left(\frac{a-b}{2}\right)$$

The equation for the resulting standing wave y_s is also shown in Equation 1. The standing wave equation can be analyzed as the product of two factors. First, consider the cosine function $\cos(\omega t)$. This expression describes simple harmonic motion over

equation 1



Standing wave equation

time; each moving particle in a standing wave oscillates in SHM with an angular frequency ω.

The cosine factor is multiplied by $2A\sin(kx)$, an expression that depends on the position x but does **not** change with time. The maximum displacement, the amplitude of the SHM, at each position x is determined by the factor $2A\sin(kx)$. This term determines an "envelope" for the particle at x; it cannot move farther from equilibrium than the absolute value of $2A\sin(kx)$ there.

Where the nodes and antinodes occur is determined by the factor $2A\sin(kx)$. If the argument to the sine function, kx, equals zero or any multiple of π, the sine is zero and there is a node at that position. If kx is any odd multiple of $\pi/2$, the sine has its maximum absolute value of one, and the wave has an antinode at that position. Because $k = 2\pi/\lambda$, two adjacent nodes or antinodes are separated by one-half a wavelength. We state equations for determining the positions of nodes and antinodes as a function of wavelength in Equation 2.

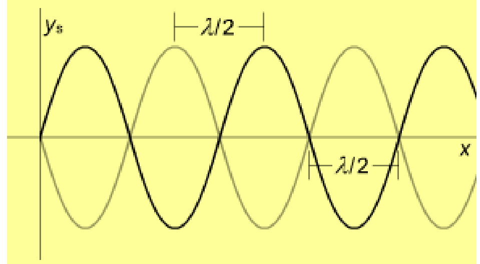
When these peaks and troughs occur is determined by the cosine function, which is a function of time. All particles reach their maximum displacement from equilibrium when the cosine equals 1 or −1, and all particles are at equilibrium when the cosine factor equals zero.

$y_1 = A\sin(kx + \omega t)$
$y_2 = A\sin(kx - \omega t)$
$y_s = y_1 + y_2$

$$\boxed{y_s = 2A\sin(kx)\cos(\omega t)}$$

y = displacement
A = amplitude
k = angular wave number
ω = angular frequency
x = position, t = time

equation 2



Nodes and antinodes

Nodes occur at

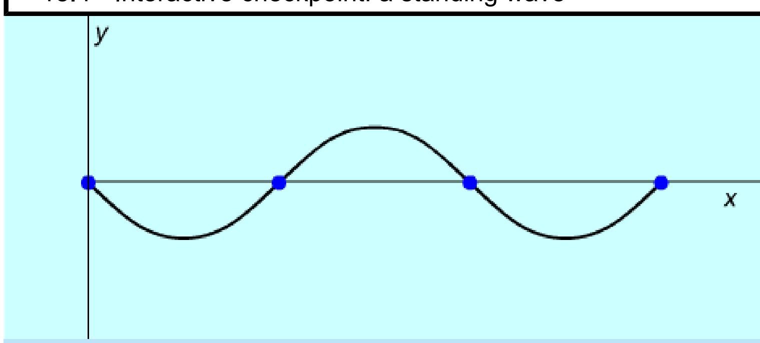
$$\boxed{x = n\frac{\lambda}{2}}$$

Antinodes occur at

$$\boxed{x = \left[n + \frac{1}{2}\right]\frac{\lambda}{2}}$$

for n = 0, 1, 2, …
λ = wavelength

18.4 - Interactive checkpoint: a standing wave



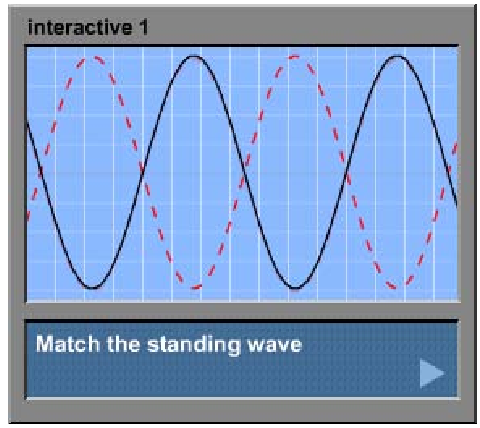
A 0.780 meter string fixed at both ends has a linear mass density and tension such that the speed of a wave in it is 95.0 m/s. If there is a standing wave in the string with 4 nodes (including the ones at the fixed ends), what is the frequency and wavelength of the standing wave?

Answer:

λ = ______ m

f = ______ Hz

18.5 - Interactive problem: match a standing wave

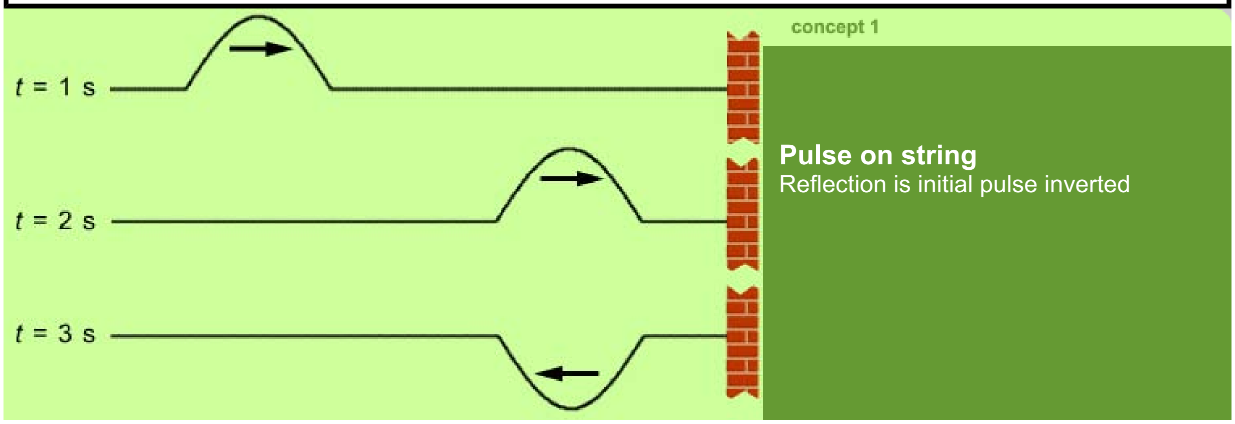


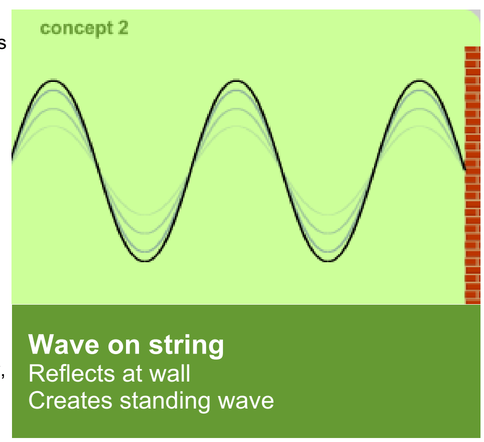
An outline, or envelope, of a standing wave on a string is shown to the right. Your goal is to match this standing wave by controlling the settings for two waves traveling in opposite directions from the ends of the string.

In the simulation, you set the amplitude and angular wave number of the traveling waves. In order to set them correctly, first determine the amplitude and wavelength of the standing wave by examining the drawing of its outline. You may find it easiest to determine the wavelength by noting points where the wave intersects the x axis. Every grid mark on the graph measures 0.50 m. From these values, and the equations in the prior section, calculate the amplitude to the nearest 0.1 m and the angular wave number to the nearest 0.1 rad/m for the two traveling waves. The angular wave number k and the wavelength λ are related by the equation $k = 2\pi/\lambda$. Set the values in the input gauges provided and press GO to see the waves combine and to test your calculations. Press RESET to start over.

If you have any trouble matching the standing wave envelope provided in the simulation, review the standing wave equations in the previous section.

18.6 - Reflected waves and resonance





We have considered standing waves formed by two waves generated by separate sources, but they can also be formed by a single wave reflecting off a fixed point. This is the basis behind the sound production of many musical instruments. We will discuss this topic using the example of a piece of string, with one end connected to a wave-making machine that vibrates sinusoidally, moving the string up and down, and the other end fixed to a wall.

To start, consider what happens when the wave machine generates a single pulse, as illustrated above. We show the pulse at three positions over time. When the pulse reaches the wall, it "yanks" on the wall. Newton's third law dictates that there will be an equal "yank" in the opposite direction, which sends a reflected pulse back down the string. The reflected pulse is inverted from the original, so when a peak reaches a wall, a trough returns in the opposite direction.

A wave is a continuous series of pulses. When the wave machine vibrates continuously, each pulse will reflect off the wall, resulting in an inverted wave moving at the same speed in the direction opposite to the original wave. The reflected wave travels down the string toward the wave machine. We also consider the wave machine as fixed, which is reasonable if its vibrations are small in amplitude. The wave then reflects off that piece of machinery just as it reflected off the wall.

The wave machine continues to vibrate, sending wave pulses down the string. If it vibrates in synchronization with the reflected wave, the result will be two traveling waves of equal amplitude, frequency and speed, moving in opposite directions on the same string. This system has created the condition for a standing wave: two identical traveling waves moving in opposite directions. You see this illustrated in Concept 2.

On the other hand, if the wave machine is not in sync with the reflected wave, it will continue to add "new" waves to the string that will combine in more complicated ways, and there may be no obvious pattern of movement on the string.

If the wave machine works in synchronization with the reflected waves to create a standing wave, we say that it is working in *resonance*. Its motion reinforces the waves, and the amplitude of the resulting standing wave will be greater than the amplitude of the vibrations of the wave machine. This is akin to you pushing a friend on a swing. If you time the frequency of your "pushes" correctly, you will send your friend higher. We will discuss next how this frequency is determined.

Fundamental frequency: The frequency of a standing wave in a vibrating string that has two nodes.

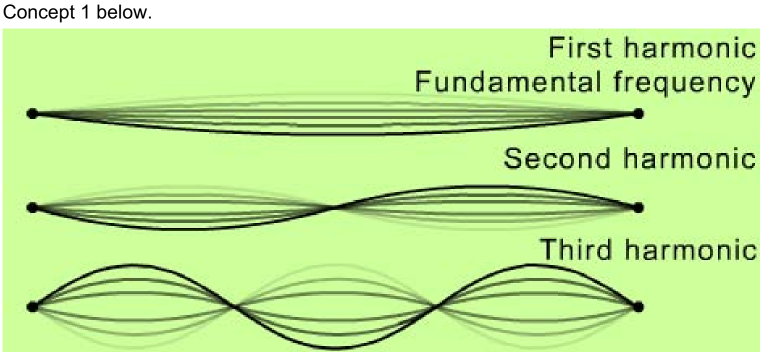
Harmonic: A frequency of a standing wave in a string that has more than two nodes.

A tuning peg is used to change the frequency of a guitar string.

We have considered vibrating strings fairly abstractly. However, strings connected to two fixed points are the basis for musical instruments such as violins, cellos and so forth. To put it another way: Orchestras have "string" sections.

In this section, we want to put your knowledge of standing waves into practice. To do so, we ask a question: When a musician plays a note, what determines the frequency at which the string will vibrate? To put the question another way, what are the possible frequencies of a standing wave on a string fixed at both ends?

To answer this question, we begin by considering the number of nodes that must be present on the vibrating string. The minimum number of nodes is two: There must be a node at each end of the string because the string is attached to two fixed points. These nodes might be the only two, but there may be intermediate nodes as well. Drawings of standing waves with zero, one and two intermediate nodes are shown in Concept 1 below.



concept 1

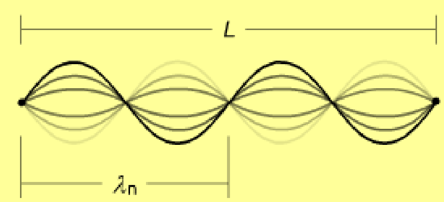
Harmonics
Standing waves at specific wavelengths
Corresponding frequencies are harmonics
First harmonic is fundamental frequency

The fundamental frequency of the string occurs when the only nodes are at the ends of the string. The fundamental frequency is also called the *first harmonic*. The second harmonic has one additional node in the middle of the string, the third harmonic two such nodes, the fourth harmonic three such nodes, and so on. Each harmonic is created by a particular *mode of vibration* of the string.

The fundamental frequency occurs when the wavelength of the standing wave is twice the length of the string, because two adjacent nodes represent half a wave. In general, the wavelength of the nth harmonic is twice the length of the string divided by n, with n being a positive integer. That is, the wavelength of the nth harmonic on a string of length L is $\lambda_n = 2L/n$.

The frequency of a wave is its speed divided by the wavelength, which lets us restate the equation above in terms of frequency. The equation to the right is the basic equation for harmonic frequencies. The harmonic frequencies are positive integer multiples of the fundamental frequency $v/2L$. Let's say the fundamental frequency f_1 of a string is 30 hertz (Hz). The second harmonic f_2 will be 60 Hz, the third harmonic f_3 90 Hz, and so forth.

equation 1



Harmonics

$$f_n = \frac{v}{\lambda_n} = n\frac{v}{2L}$$

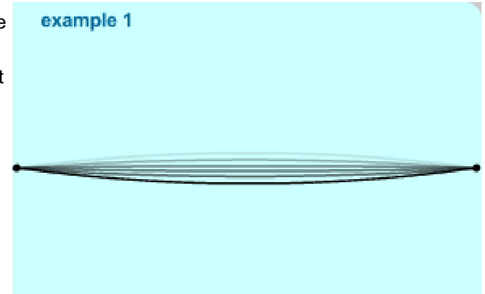
for $n = 1, 2, 3, \ldots$
f_n = n^{th} harmonic, v = wave speed
λ_n = wavelength of n^{th} harmonic
L = string length

You see these principles in play in musical instruments like the piano. Its strings are of different lengths, which is one factor in determining their fundamental frequency. Other factors that you see in the equations are also employed in musical instruments to determine a string's fundamental frequency. Pianos have thicker and thinner strings. A string's mass per unit length (its linear density) will partly determine its fundamental frequency, by changing the wave speed on the string.

In addition, string instruments are "tuned" by changing the tension in a string. You will see a guitarist frequently adjusting her instrument by turning a tuning peg, as you see at the top of this page, which determines the tension in a string. Along with linear density, this is the other factor that determines wave speed. A guitarist will also press her finger on a single string to temporarily create a string with a specific length and fundamental frequency.

A harmonic is sometimes called a *resonance frequency* or a *natural frequency*. Musicians also use other terms dealing with frequencies and harmonics. An *overtone* or a *partial* is any frequency produced by a musical device that is higher than the fundamental frequency. Unlike a harmonic, the frequency of an overtone does not necessarily bear any simple numerical relationship to the fundamental.

Some overtones are harmonics − that is, whole-number multiples of the fundamental − but some are not. Musical instruments such as drums, whose modes of vibration can be very complex, create non-harmonic overtones. Although harmonic overtones are "simple" integer multiples of the fundamental, they are often referred to by numbers that are, confusingly, "one off" from the numbers for harmonics: The first overtone is the second harmonic, and so on. Each musical instrument has a characteristic set of overtones that creates its distinctive timbre.



A tension of 417 N is applied to a 1.56 m string of mass 0.00133 kg. What is the fundamental frequency?

$$v = \sqrt{\frac{F}{m/L}}$$

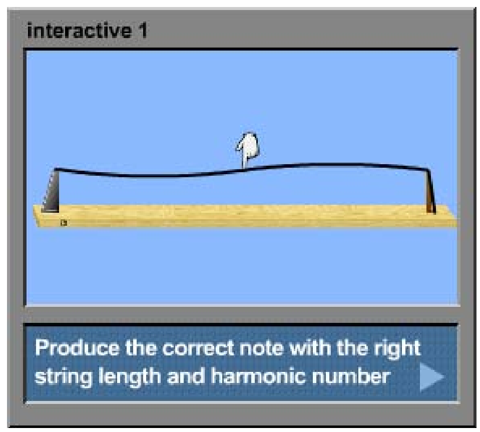
$$v = \sqrt{\frac{417\ \text{N}}{(0.00133\ \text{kg})/(1.56\ \text{m})}}$$

v = 699 m/s

$$f_n = n\frac{v}{2L}$$

f_1 = (1)(699 m/s)/2(1.56 m)
f_1 = 224 Hz

18.8 - Interactive problem: tune the string



Let's say you play an unusual instrument in the orchestra, the alto pluck, an instrument specifically designed for physics students. The concert is underway and your big moment is coming up, when you get to play a particular note on the pluck.

You are supposed to play the G above middle C, a note that has a frequency of 392 Hz. You produce the correct note by setting the string length and the harmonic number. Remember that harmonic numbers are multiples of the fundamental frequency of the string. For this instrument, you are limited to harmonic numbers in the range of one to four.

When you set a harmonic number higher than one, a finger will touch the string at a position that will cause the string to vibrate with the harmonic number you selected. It does so by creating a standing wave node at the appropriate location. For instance, if you choose a harmonic number of four, there will be three nodes between the two ends of the string, and the finger will be one-fourth of the way along the string. If you see a musician such as a cellist performing, you will see that he sometimes lightly places his fingers at locations along a string to create harmonics in just this way. He also frequently presses a string firmly against the fingerboard to create a different, shorter, string length.

The string length of the alto pluck can range from 1.00 to 2.00 meters, but within that range, you are restricted to values that will produce frequencies found on the musical scale. You will find that as you click on the arrows for length, the values will move between appropriate string lengths. The harmonic values are easy to set: Just choose a number from one to four!

The description above is reasonably complicated; it is impressive that musicians with some instruments must make similar determinations as they play. In terms of approaching this problem, you will want to start with the equation

$$f_n = nv/2L$$

that enables you to calculate the frequency of the nth harmonic. The wave speed in your stringed instrument is 588 m/s.

There is only one solution to the problem we pose, given the range of string lengths and harmonics together with the wave speed we provide.

If you are not sure how to proceed, try solving the equation above for the string length and entering the known values. You will still have another variable, n, left in the equation. However, since you know the range of string lengths, and that n must be an integer from one to four, you will be able to solve the problem.

To test your answer, set a string length and harmonic number and press GO to see a hand come down and pluck the string so you can hear the resulting note and see its frequency. The resulting musical note is displayed on the sounding board. Press RESET to start over.

18.9 - Wave reflecting at a movable point

concept 1

Reflection at a movable point
Pulse is reflected, but not inverted

Here, we again consider a wave reflected on a string. In this case, though, one end of the string is attached to a ring that can slide up and down freely, instead of being attached to a fixed point. As before, a device that can generate a pulse or wave is connected to the other end. The pulse or wave will reflect when it reaches the sliding ring, but unlike the reflection of a string attached to a fixed point, this kind of reflection does not cause a reversal of the displacement of the pulse or wave. A peak reflects as a peak, and a trough reflects as a trough. In an ideal movable end-point configuration, as with a fixed end-point, neither the amplitude nor the frequency of the reflected wave changes.

Although you may wonder a bit why one would attach a string to a sliding ring and watch reflection patterns, there are good reasons for such a thought experiment. This configuration provides the model for explaining standing waves created in some wind instruments (vibrating air columns, where the pipe is open to atmosphere), as well as certain instances of light reflection.

18.10 - Music from wind instruments

Transverse waves often serve as the example of standing waves since the nodes and antinodes can be easily depicted. However, if you have listened to a flute or an organ, or other wind instruments, you have enjoyed music produced by longitudinal standing waves.

We use a simplified organ pipe, shown in Concept 1, as our setting for a standing longitudinal wave. The pipe is open at both ends. The particles that make up the standing wave inside the pipe oscillate back and forth in simple harmonic motion. That is depicted with the waveforms you see below the pipe, which represent the particle displacement from equilibrium.

Why is there a standing wave? Like a wave on a string, the sound wave that creates the standing wave reflects at each end of a pipe. A wave reflecting off an “opening” may seem counterintuitive, but consider the nature of sound waves: They are alternating regions of higher and lower pressure. The area outside a pipe remains at a constant atmospheric pressure, acting like a wall to the wave. Not all of the wave reflects back, but the portion that does is enough to create a standing wave.

Displacement antinodes are located (approximately) at the ends of the open pipe, where the particles move back and forth with the greatest displacement. The opposite occurs in stringed instruments, where there are nodes at the ends of a string where it does not move. In a longitudinal wave, the areas of greatest displacement have the lowest pressure, so the displacement **antinodes** are also pressure **nodes**.

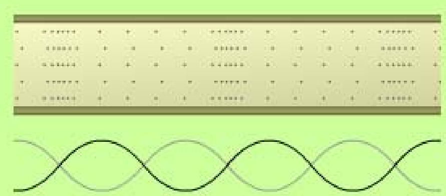
The harmonics of a pipe that is open on both ends can be calculated using the same equation as for a standing wave on a string, which is shown to the right. The first harmonic, the fundamental frequency, equals the wave speed divided by twice the length of the pipe. The other harmonics are multiples of that value.

If the pipe has one closed end, that end acts as a displacement node. A clarinet is an example of an instrument that behaves like this, since the end with the mouthpiece is essentially closed. The particles do not move there because the closed end prevents them from doing so. In an instrument like this, only the odd harmonics (first, third, fifth, and so on) are produced. This is expressed in Equation 2. Notice that this equation has a denominator of four times the pipe length, not two. Why? Because the distance between a node and an adjacent antinode is one-fourth of a wavelength.

The different nodes at open and closed pipe ends correspond to the different types of reflection that occur when a traveling wave on a string reflects off of a fixed or a sliding point.

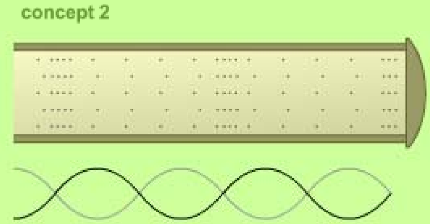
You can see physical expressions of these equations in various musical instruments. Consider an organ: It consists of pipes of different lengths, which have different fundamental frequencies. Or you can think about a trombone. It has a movable part that allows the musician to alter the length of an air column and change its fundamental

concept 1



Standing waves in pipes
Waves reflect at ends of pipe
Create standing wave
Open end is displacement antinode

concept 2



At closed end
Closed end is displacement node

equation 1

Harmonics in pipe open at both ends

frequency.

The speed of a wave in a string depends on both its tension and linear density, and a musician can tune at least the tension to change the wave speed and harmonics of her stringed instrument. The speed of a sound wave in a pipe depends on the bulk modulus and density of the fluid in the pipe, and these are harder to change! But there is one way to change the speed of sound and “tune” the harmonics of a particular wind instrument: the human throat. Inhaling helium, which has a lower density and a higher speed of sound (965 m/s) than air (343 m/s), gives a person a high-pitched cartoonish voice.

$$f_n = \frac{nv}{2L}$$

for $n = 1, 2, 3, \ldots$
f_n = n^{th} harmonic
v = wave speed
L = pipe length

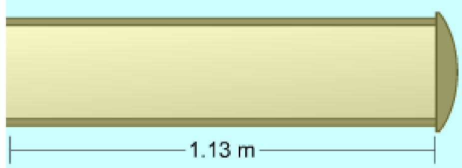
equation 2

Harmonics in pipe with one closed end

$$f_n = \frac{nv}{4L}$$

for $n = 1, 3, 5, \ldots$
f_n = n^{th} harmonic
v = wave speed
L = pipe length

example 1



What is the pipe’s lowest harmonic above its fundamental frequency?

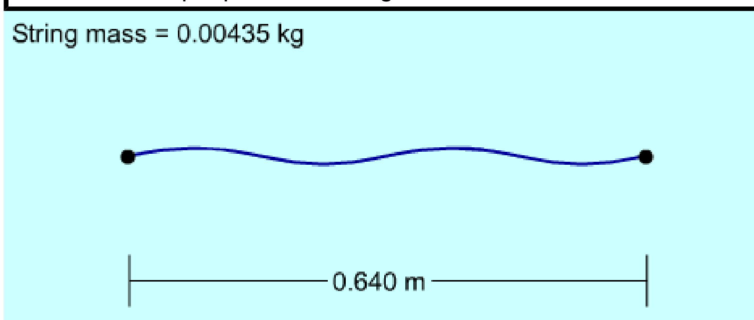
$$f_n = \frac{nv}{4L}$$

$n = 3$, $v = 343$ m/s

$$f_3 = \frac{3\,(343\text{ m/s})}{4\,(1.13\text{ m})}$$

$f_3 = 228$ Hz

18.11 - Sample problem: string tension



The frequency of the standing wave on the string is 329 Hz. What is the tension on the string?

The values shown in the problem are representative of the lowest frequency string on a guitar. To answer the question, you must first determine the harmonic of the standing wave. You can do so by inspecting the diagram above. We exaggerated the amplitude of the wave to make the nodes and antinodes more visible.

Variables

string length	L = 0.640 m
string mass	m = 0.00435 kg
frequency	f = 329 Hz
harmonic number	n
wave speed	v
tension	T

What is the strategy?

1. Determine the harmonic by counting the number of nodes shown in the string above.
2. Calculate the wave speed from the frequency, string length and harmonic number.
3. Calculate the tension using the equation for wave speed on a string.

Physics principles and equations

Nodes are locations where there is no displacement. The harmonic number is one less than the number of nodes, including the nodes at the ends of the string.

The equation for the nth harmonic

$$f_n = \frac{nv}{2L}$$

The wave speed on a stretched string

$$v = \sqrt{\frac{T}{m/L}}$$

Step-by-step solution

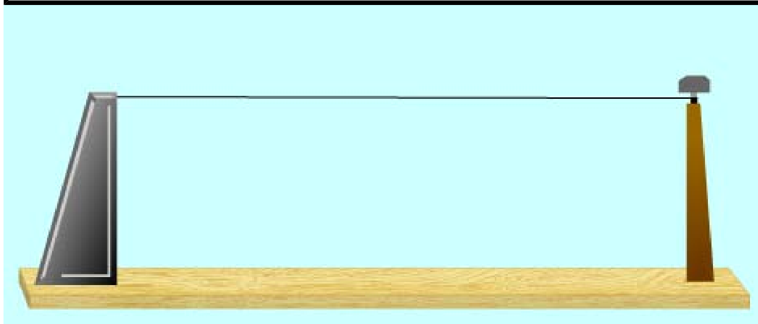
First, we determine the harmonic number by looking at the wave above. Then, we calculate the wave speed using the equation for the frequency of the nth harmonic.

Step	Reason
1. $n = 4$	harmonic is one less than number of nodes
2. $f_n = \frac{nv}{2L}$	frequency of nth harmonic
3. $v = \frac{2Lf_n}{n}$	solve for wave speed
4. $v = \frac{2(0.640\ \text{m})(329\ \text{Hz})}{4}$	substitute values
5. v = 105 m/s	evaluate

Now we use the equation that relates wave speed to tension, mass, and string length.

Step	Reason
6. $v = \sqrt{\frac{T}{m/L}}$	wave speed on string
7. $T = \frac{v^2 m}{L}$	solve for tension
8. $T = \frac{(105\ \text{m/s})^2(0.00435\ \text{kg})}{0.640\ \text{m}}$	substitute values
9. T = 74.9 N	evaluate

18.12 - Interactive checkpoint: tune the string

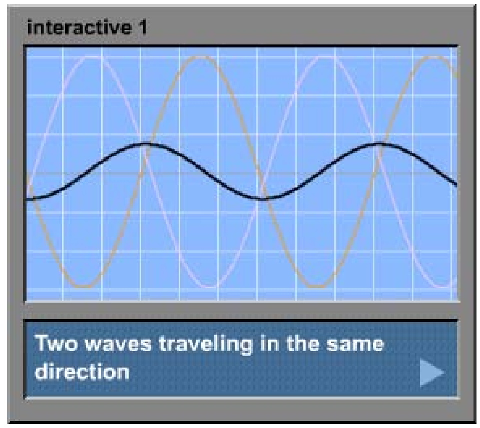


A string of length 0.300 m and mass 0.00870 kg can be tuned to any tension from 355 N to 555 N. What should the tension be for the string to produce a tone of frequency 587 Hz? Hint: This note is not the fundamental frequency; consider the harmonic frequencies of the string.

Answer:

$T =$ ______ N

18.13 - Interactive problem: waves traveling in the same direction



We have shown how waves combine when they are moving in opposite directions in a medium. In this section, you can observe what happens with traveling waves moving in the **same** direction in a medium.

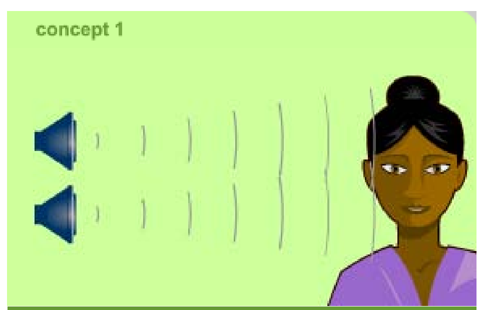
The interactive simulation on the right shows two waves traveling in the same direction in the same medium, and the resulting combined wave. The two waves that combine are shown in light orange and pink, and the combined wave is black.

The individual waves have the same amplitude, wavelength and speed. However, they can have different phase constants, which means they can have a phase difference. You can set the phase difference.

Can you determine in advance a phase difference that will cause a combined wave with the greatest amplitude? A phase difference that will cause a "flat" wave? In essence, the simulation is adding two sine functions, so consider the phase difference that will create two identical waves, or two waves with opposite displacements.

Also note an important point: is the combined wave a standing wave? Or does it travel?

18.14 - Wave interference and path length



Path lengths the same
Constructive interference

We described wave interference resulting from a phase difference, or differing initial conditions for two waves moving in the same direction. Interference also results when two waves travel from different starting points and meet. To illustrate this, we use the example of two longitudinal traveling waves produced by two loudspeakers.

To the right, we show a person listening to the loudspeakers. The vibrating speakers create regions of pressure that are greater than atmospheric pressure (condensation) and less (rarefaction). We show this pattern of oscillation emanating from each speaker in Concept 1. We assume the speakers create waves with the same amplitude and wavelength, and that there is no phase difference between them. We focus on the point where the two waves combine just as they reach the listener's ear.

In Concept 1, we position the two speakers so that the listener is equidistant from them. The distance from a speaker to the ear is called the *path length*. Since the waves travel the same distance, they will be in phase when they arrive. This means peaks and troughs exactly coincide with each other.

When two peaks combine, they double the pressure increase. When two troughs combine, they also add, and the pressure decrease is doubled. The listener hears a louder sound. In short, the waves constructively interfere.

The waves would also constructively interfere if one speaker were placed one wavelength farther away. Peaks would still meet peaks and troughs would still meet troughs. In fact, if the difference in the distances between the loudspeakers and the listener is any integer multiple of the wavelength, the waves constructively interfere. (This does assume that the loudspeakers vibrate at a constant frequency, an unusual assumption for most music.) You see this condition for constructive interference stated in Equation 1.

Can we arrange the speakers so that the waves **destructively** interfere? Yes, by moving one speaker one-half wavelength away from the listener. This is shown in Concept 2. When the waves combine, peaks will meet troughs and vice versa. This will also occur if the speaker/listener distances differ by half a wavelength, or 1.5 wavelengths, or any half-integer multiple of the wavelength. This condition for destructive interference is stated as an equation on the right. If the waves have the same amplitude at the ear and destructively interfere completely, the result is silence at the listener's position.

The principle of conservation of energy applies to sound waves. When two waves destructively interfere at some position, the energy does not disappear. Rather, there must be another area where the waves are constructively interfering, as well. For any area of silence, there must be a louder area. Positioning loudspeakers or designing concert halls to minimize the severity of these "dead spots" and "hot spots" is a topic of interest to audiophiles and sound engineers alike.

concept 2

Path lengths differ by one-half wavelength
Destructive interference

equation 1

Path lengths differ by one-half wavelength

Constructive interference:

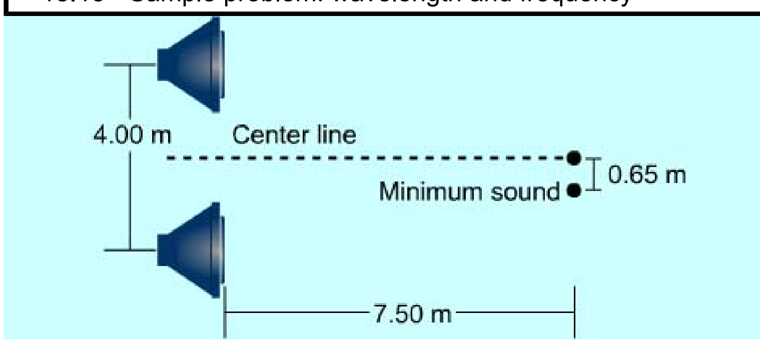
$$\Delta p = n\lambda$$

Destructive interference:

$$\Delta p = (n+\tfrac{1}{2})\lambda = \frac{2n+1}{2}\lambda$$

Δp = difference in path lengths
$n = 0, \pm 1, \pm 2, \ldots$
λ = wavelength

18.15 - Sample problem: wavelength and frequency

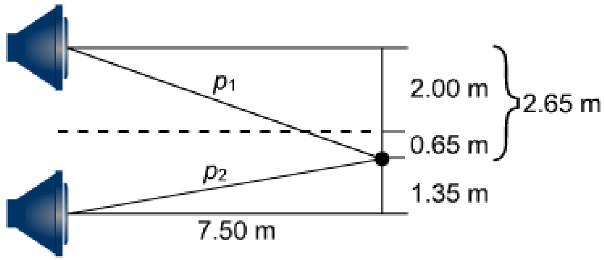


Two loudspeakers emit identical sound waves. A listener at the center position hears the loudest sound. At a point 0.65 m away from the center the listener hears the least sound. What is the sound's wavelength and frequency?

The sound intensity reaches a minimum when destructive interference occurs.

Draw a diagram

We need to know the path lengths from the speakers to the listener when the sound is at a minimum. The hypotenuses of the two triangles drawn here are the distances we want, labeled p_1 and p_2.



Variables

To calculate the frequency, we need the speed of sound in air, shown below.

path length from top speaker	p_1
path length from bottom speaker	p_2
wavelength	λ
frequency	f
speed of sound $(20°\mathrm{C})$	$v = 343$ m/s

What is the strategy?

1. Calculate the two path lengths when the listener is at the minimum sound position.
2. Use the difference in path lengths and the fact that destructive interference is occurring to calculate the wavelength.
3. Use the wavelength and the speed of sound to calculate the frequency.

Physics principles and equations

Destructive interference occurs when the difference in path length is an odd multiple of half the wavelength.

$$\Delta p = \frac{2n+1}{2}\lambda$$

Since the listener is at the **first** position where a minimum occurs, the path lengths differ as little as possible for destructive interference. This means $n = 0$ and the equation becomes

$$\Delta p = \lambda/2$$

The wavelength and frequency for a wave with speed v are related by the equation

$$f = v/\lambda$$

Mathematics principles

We will use the Pythagorean theorem for a right triangle with legs a and b and hypotenuse c, here stated to solve for the hypotenuse.

$$c = \sqrt{a^2 + b^2}$$

Step-by-step solution

We start by calculating the two path lengths, and subtracting to calculate their difference.

Step	Reason
1. $\Delta p = p_1 - p_2$	difference in path lengths
2. $p_1 = \sqrt{(2.65\text{ m})^2 + (7.50\text{ m})^2}$	Pythagorean theorem
3. $p_1 = 7.95$ m	evaluate
4. $p_2 = \sqrt{(1.35\text{ m})^2 + (7.50\text{ m})^2}$	Pythagorean theorem
5. $p_2 = 7.62$ m	evaluate
6. $\Delta p = 0.33$ m	substitute equations 3 and 5 into equation 1

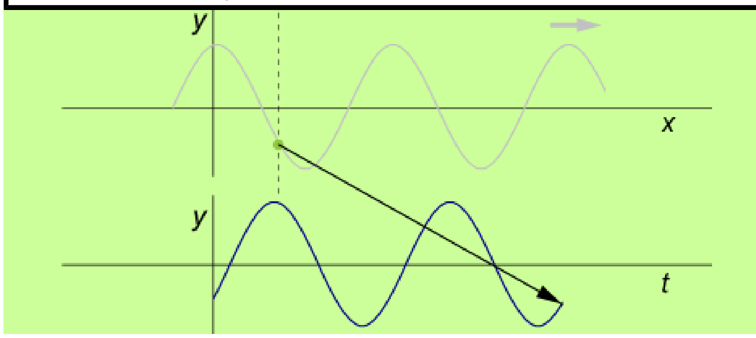
The listener is at the first position where destructive interference occurs, so the paths differ by exactly one-half the wavelength of the sound.

Step	Reason
7. $\Delta p = \lambda/2$	destructive interference
8. $(0.33\text{ m}) = \lambda/2$	substitute equation 6 into equation 7
9. $\lambda = 0.66\text{ m}$	solve and evaluate

From the wavelength and the speed of sound, we can calculate the frequency.

Step	Reason
10. $f = v/\lambda$	frequency-wavelength equation
11. $f = (343\text{ m/s})/(0.66\text{ m})$	substitute values
12. $f = 520\text{ Hz}$	divide

18.16 - Traveling wave viewed at a fixed position



concept 1

Displacement at a fixed *x* position

Particle moves vertically in SHM

· As graph of displacement over time shows

One way to represent a traveling wave, as shown in the top graph in the illustration above, is to graph the y displacement versus the x position, and let the graph change over time. This is the representation of traveling waves we have used so far in this chapter. The arrow above the graph indicates its direction of motion over time.

Another way to look at a traveling wave is to pick a particular fixed x position and consider the y displacement of a particle at that position over time. Since the traveling wave is defined by a sinusoidal function of both position and time, the displacement at the fixed position is a sinusoidal function of time. This means the displacement of a particle at any fixed x position on a traveling wave exhibits simple harmonic motion (SHM).

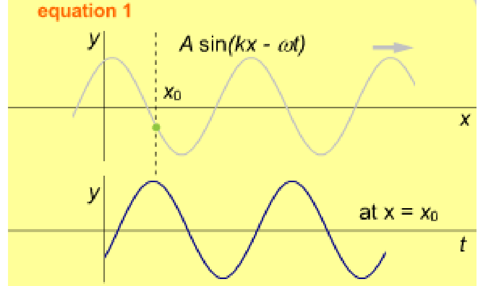
You can see this in the illustration above. The top graph is a snapshot in time of the traveling wave, with the y displacement graphed against the x position. A particular x position is highlighted. Imagine the y displacement at that position over time: It would vary sinusoidally as the traveling wave moved past. We show the y displacement at this position "distributed" over time on the bottom graph above. Unlike the first graph, the horizontal axis here is time, not x position.

In Equation 1, we show the equation for the y displacement caused by a traveling wave at a particular fixed x position over time. It is the standard equation for SHM. The phase constant φ depends on the x position chosen. For example, at the position $x = 0$, the equation for the y displacement is

$$y = A\cos(\omega t + \pi/2)$$

Regardless of the x position you choose, the y displacement obeys simple harmonic motion.

equation 1



Displacement at a fixed *x* position

At any fixed position:

$$y = A\cos(\omega t + \varphi)$$

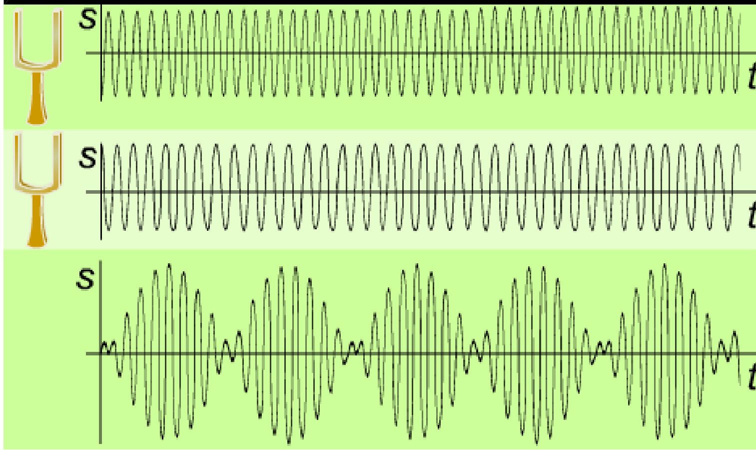
y = displacement
A = amplitude
ω = angular frequency
t = time, φ = phase constant

We derive the equation for the y displacement at a fixed x position below, using the traveling wave equation $y = A\sin(kx - \omega t)$ together with the trigonometric identity $\sin(\theta) = \cos(\pi/2 - \theta)$. Let the fixed x position be $x = x_0$.

Step	Reason
1. $y = A\sin(kx_0 - \omega t)$	y displacement at $x = x_0$
2. $y = A\cos(\pi/2 - kx_0 + \omega t)$	trigonometric identity
3. $y = A\cos(\omega t + \varphi)$	$\varphi = \pi/2 - kx_0$

The phase constant, then, is $\varphi = \pi/2 - kx_0$. When the fixed point x_0 is zero, then $\varphi = \pi/2$, as claimed above.

18.17 - Beats



concept 1

Beats
Periodic variations in amplitude produced by two waves of different frequencies

Beats: The pattern of loud/soft sounds caused by two sound sources with similar but not identical frequencies.

So far in this chapter, we have studied combinations of traveling waves with the same amplitude, frequency and wavelength. You may wonder what happens if waves that have different properties are combined, and in this section, we consider such a situation.

To be specific, we consider two traveling waves with the same amplitude but slightly different frequencies, using sound waves as our example. We examine the waves and their combination at a fixed x position. You see the graphs above of two such waves over time **at a particular position**. Note that in this case the graphs that you see are displaying the displacement over time of a particle at a fixed position in a medium carrying longitudinal waves.

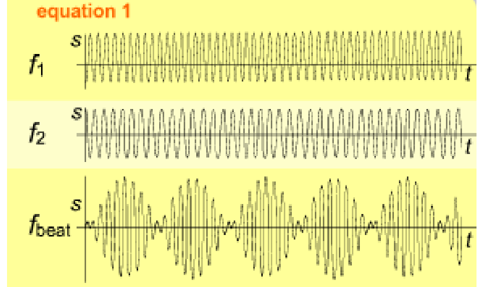
The waves were created by tuning forks, and the combined wave is shown below them. When the waves combine, they produce a wave whose amplitude is not constant, but instead varies in a repetitive pattern. For sound waves, these "beats" are heard as a repeating pattern of variation in loudness in a wave of constant frequency.

Musicians sometimes use beats to tune their instruments. Sounds that are close in frequency produce audible beats, but the beats disappear when the frequencies are the same. A guitar player, for example, might tune the "A" string by playing an "A" on another string at the same time and adjusting the tension of the "A" string until there are no longer any audible beats. This occurs when the frequencies match exactly, or are close enough that the beats are so far apart in time they can no longer be heard.

We hear beats because the waves constructively and destructively interfere over time. When they constructively interfere, there is greater condensation and rarefaction of the air at our ears and we hear louder sounds. Destructive interference means a smaller change in pressure and a softer sound.

The *beat frequency* equals the number of times per second we hear a cycle of loud and soft. This is computed as shown in Equation 1, as the difference of the original frequencies. When the frequencies are the same, the beat frequency is zero, as when two strings are perfectly in tune. Humans can hear beats in sound waves at frequencies up to around 20 beats per second. Above that frequency, the beats are not distinguishable.

equation 1

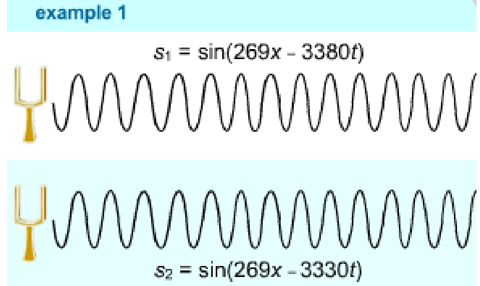


Beat frequency

$$f_{beat} = f_1 - f_2$$

f_{beat} = beat frequency
f_1 = frequency of sound one
f_2 = frequency of sound two
$(f_1 > f_2)$

example 1



What is the beat frequency when these two waves combine?

$f = \omega/2\pi$
$f_1 = (3380 \text{ rad/s})/2\pi = 538 \text{ Hz}$
$f_2 = (3330 \text{ rad/s})/2\pi = 530 \text{ Hz}$
$f_{beat} = f_1 - f_2$
$f_{beat} = 538 \text{ Hz} - 530 \text{ Hz}$
$f_{beat} = 8 \text{ Hz}$

Fourier analysis: A technique for expressing any periodic function as the sum of a series of sinusoidal functions.

Complex musical waveforms.

Fourier analysis is an important tool in many engineering, scientific and mathematical applications. Here, we provide an introduction to the topic. Fourier analysis is a method of expressing any periodic function as a sum of a series of sinusoidal functions. It is a technique widely used in computer based audio compression. It can also be used to analyze the waveforms produced by musical instruments.

concept 1

$s(t) = \sin(\omega t) + 0.3\sin(3\omega t) + 0.3\sin(5\omega t) + 0.4\sin(7\omega t)$

Fourier analysis
A complex periodic function can be created as sum of series of sine waves

Above, you see the sound waveforms for a clarinet and a trombone, both of which are playing a B-flat at the same frequency of 233 Hz, plotted against time at a fixed location in space − say, at the position of a listener's ear. These waves may look complicated, but each is composed of the fundamental frequency and harmonics, combined in varying intensities. It is these varying harmonic intensities that make musical instruments sound different.

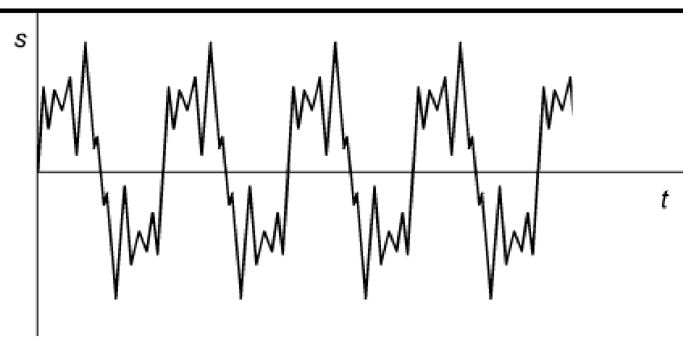
A tuning fork, for example, has a strong fundamental frequency but very little intensity in other harmonics. Its waveform is close to being purely sinusoidal. A harmonica, on the other hand, has significant contributions from many harmonic frequencies, giving it a very different sound. The combination of harmonic frequencies in varying amounts creates the different waveforms of musical instruments.

A periodic function is one that repeats consistently over time. If $F(t) = F(t + T)$ for every value of t, then the function F has the period T. Any periodic function can be expressed as the sum of sinusoidal waves in a harmonic series. The French mathematician Jean Baptiste-Joseph Fourier made this remarkable discovery in the early 1800s. Fourier analysis uses a series of sinusoidal waves added together to approximate periodic functions. The more wave functions that are added, the closer the approximation to the original wave.

In Concept 1 to the right, we show the same clarinet wave as above and a sum of sine functions that are added together to approximate it. The shape of a clarinet is close to that of a pipe closed at one end, which produces only odd harmonic standing waves. So the first, third, fifth and seventh harmonics are added to approximate the clarinet wave. Choosing the coefficients for the harmonics correctly is the key to Fourier analysis. There are mathematical techniques for doing this, but we will not discuss them here.

Fourier analysis is used in a variety of ways, including when information needs to be stored and used in digital form. Compact disc technology is based on the techniques of Fourier analysis. Many music synthesizers create complex waveforms by adding simpler component waves. In audio compression, the coefficients of the sine functions approximating a sound wave can be saved, instead of recording every point on the wave, and the wave can then be recreated from this smaller amount of information. To cite another example, geophysicists use Fourier analysis to study seismic waves. These are just a few of the many applications of this powerful tool.

18.19 - Interactive problem: Fourier synthesis



A periodic wave like this can be created by summing harmonics.

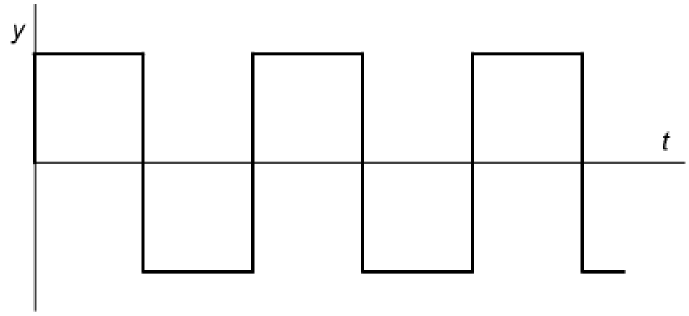
Above, you see the graph of a periodic wave. This is a graph of the wave's displacement over time at a particular position. The wave may be transverse or longitudinal. By means of Fourier analysis, any such periodic wave can be expressed as the sum of a series of sinusoidal functions. The functions are all harmonics of some fundamental frequency. Finding the amplitudes of the harmonic functions is the key to Fourier analysis.

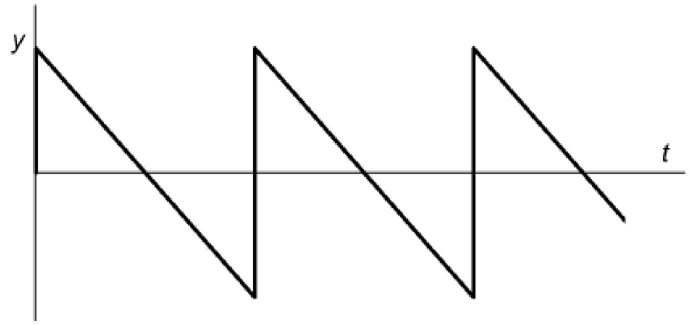
The reverse process − adding harmonic functions with different amplitudes to create a new periodic wave − is called Fourier synthesis. Although the mathematics of Fourier analysis and synthesis are beyond the scope of this textbook, you can use the simulation to the right to try your hand at synthesizing waves by adding harmonics of a fundamental frequency.

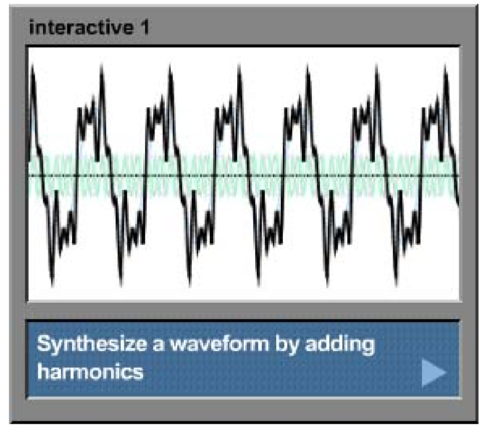
You can add up to three harmonic waves to a fundamental wave. For each of the harmonics you add, you can choose the harmonic number and the amplitude. The fundamental wave has amplitude 1.0 nanometers (1.0×10^{-9} m), and the harmonics have amplitudes no greater than this. (Note: This is a typical displacement amplitude for a sound wave in air.)

You cannot change the amplitude of the fundamental frequency. The amplitudes of the harmonics are initially set to 0.0 nanometers, and all are initially set as second harmonics. There is no GO button for this simulation, just a RESET button that will return the values to their initial settings. Change the amplitudes and harmonic numbers and see what the result is for the summed wave. What is the difference when you add in “low order” harmonics (2, 3, 4) versus “high order” harmonics (8, 9, 10)? Do odd and even harmonics affect the wave differently?

If you want a goal, see how close you can get to the two waveforms below, one of which is called a square wave and the other a sawtooth wave. You will not be able to match these waveforms exactly, but you can get close.







18.20 - Gotchas

A standing wave has its maximum displacement at an antinode. Yes, that is correct. An antinode is the opposite of a node, where no motion occurs.

The locations where peaks and troughs occur are constant in a standing wave. That is correct, and this is the distinguishing point between a standing wave and a traveling wave.

I see a standing wave on a string with two fixed ends and a single antinode. The wave has two nodes. Yes. The two fixed ends are nodes.

Two waves traveling in the same direction can cause a standing wave. No. Waves traveling in opposite directions can cause a standing wave. However, a wave from a single source when it is reflected can cause a standing wave, because the reflected wave is traveling in the opposite direction.

18.21 - Summary

The principle of linear superposition says that whenever waves travel through a medium, the net displacement of the medium at any point in space and at any time is the sum of the individual wave displacements. The superposition of waves explains the phenomenon of interference.

Destructive interference occurs when two waves in the same medium cancel each other, either partially or fully. If two sinusoidal waves have the same wavelength, destructive interference will happen when the waves are close to being completely out of phase, meaning that their phase constants differ by π radians (or $180°$).

Constructive interference occurs when two waves in the same medium reinforce each other. This happens when the waves are close to being in phase for waves with the same wavelength.

Intermediate interference is a general term for a situation where the difference between the phases of two interfering waves, called the phase shift, is somewhere between 0 and π radians.

When two identical waves traveling in opposite directions in the same medium interfere, they produce a standing wave. In standing waves, there are points called nodes that experience no displacement at all. The points that experience the maximum displacement are called antinodes. Standing waves can be either transverse, as with the oscillations of a piano string, or longitudinal, as with the sound waves in an organ pipe.

Identical waves from different sources can interfere with each other at a point in space based on the distances they travel to that point, called their path lengths. Provided they start out in phase, if the path lengths of two waves to a certain point differ by an integer number of wavelengths, they will constructively interfere at that point. If the path lengths differ by a half-integer number of wavelengths, they will exhibit completely destructive interference.

Another type of interference occurs when two waves have different frequencies. In this case, beats are produced. The waves alternate between constructive and destructive interference.

A technique called Fourier analysis is used to decompose a periodic function into a series of sinusoidal functions. This can be useful for finding the individual sinusoidal waves, a fundamental plus harmonics, that make up the combined wave.

Equations

Standing wave

$y_s = 2A \sin(kx)\cos(\omega t)$

Nodes: $x = n\frac{\lambda}{2}$

Antinodes: $x = \left[n + \frac{1}{2}\right]\frac{\lambda}{2}$

Harmonics

$f_n = \frac{v}{\lambda_n} = n\frac{v}{2L}$

$f_n = \frac{nv}{2L}$ for pipe, both ends open

$f_n = \frac{nv}{4L}$ for pipe, one end closed

Interference

Constructive: $\Delta p = n\lambda$

Destructive: $\Delta p = \frac{2n+1}{2}\lambda$

Beat frequency

$f_{beat} = f_1 - f_2$

Chapter 18 Problems

Conceptual Problems

C.1 A guitar designer is trying out a string of mass per unit length μ. In order to have a fundamental frequency of f_o, the required tension in this string is F_T. However, she finds that when this tension is applied, this string snaps under the applied stress. Will using a string made of the same material, only thicker, solve her design problem? Explain your answer.

○ Yes ○ No

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the first interactive problem in this section to create a situation where the pulses will exactly cancel each other out. (a) If the left pulse is a peak, what should the right pulse be? (b) If the amplitude of the left pulse is 2.00 m, what should the amplitude of the right pulse be? (c) If the width of the left pulse is 1.00 m, what should the width of the right pulse be?

(a) ○ Peak ○ Trough
(b) __________ m
(c) __________ m

Section 3 - Standing wave equations

3.1 A standing wave in a string is described by the equation
$y_s = (5.0\times10^{-4})\sin(4.8\times10^{-3}x)\cos(1400t)$ meters
What is the x position (a) of the 0^{th} antinode? (b) of the 3^{rd} node?

(a) __________ m
(b) __________ m

3.2 The 4th node of a standing wave occurs at a position of 4.10 meters. Where is the 2nd antinode?

__________ m

3.3 A string with a linear mass density of 0.0150 kg/m and length of 0.750 m is fixed at both ends with a tension of 45.0 N. If you wish to start a standing wave in the string with 6 nodes (including the nodes at the fixed ends),
what is the (a) angular wave number and (b) the angular frequency that you will use? Hint: With 6 nodes, the last node corresponds to $n = 5$.

(a) __________ rad/m
(b) __________ rad/s

Section 5 - Interactive problem: match a standing wave

5.1 Use the information given in the interactive problem in this section to calculate (a) the amplitude and (b) the angular wave number of the wave traveling from left to right. Test your answer using the simulation.

(a) __________ m
(b) __________ rad/m

Section 7 - Harmonics

7.1 What is the fundamental frequency of a 4.65×10^{-3} kg, 2.50-meter string under 438 newtons of tension?

__________ Hz

7.2 A string that is under a tension of 389 N has a linear mass density of 0.0220 kg/m. Its fundamental frequency is 440 Hz (an A note). How long is the string?

__________ m

7.3 A string is fixed to two eye hooks 0.940 m apart. The tension in the string is 505 N and its mass is 2.30×10^{-4} kg. What is the frequency of the 3rd harmonic?

__________ Hz

7.4 You drape a string over a pulley and hang a mass of 35.0 kg from one end. You tie the other end to a point a distance L from the pulley so that the string is horizontal between this point and the pulley. The string has a linear mass density of 0.0470 kg/m. If you want the difference in frequency between any two consecutive harmonics to be 35.0 Hz, what will the distance L have to be?

__________ m

7.5 A musical instrument is constructed using a 0.550-meter wire with a mass of 6.90×10^{-4} kg. If the fundamental frequency is to be 523 Hz, what is the required tension in the wire?

__________ N

7.6 A harp string plays the A above middle C, vibrating at its fundamental frequency of 440 Hz. The tension in the string is 57.8 N, and its mass density is 5.72×10^{-4} kg/m. What is the distance between the nodes of the standing wave pattern?

__________ m

Section 8 - Interactive problem: tune the string

8.1 Use the information given in the interactive problem in this section to calculate (a) the string length and (b) the harmonic required to play a G note. Test your answer using the simulation.

(a) __________ m

(b) i. 1
ii. 2
iii. 3
iv. 4
v. 5

Section 10 - Music from wind instruments

10.1 A 2.1-meter pipe is open at both ends. What is its fundamental frequency? (Use 343 m/s for the speed of sound.)

__________ Hz

10.2 A pipe with one closed end has a fundamental frequency of 262 Hz at 42.8°C. How long is the pipe?

__________ m

10.3 A pipe with both ends open has a fundamental frequency of 492 Hz. You cover one end of the pipe. What is the new fundamental frequency?

__________ Hz

10.4 If a pipe with two open ends has a third harmonic at 308 Hz at 20.0°C, (a) what is the fundamental frequency? (b) What is the temperature when the fundamental frequency is 99.9 Hz for the same pipe?

(a) __________ Hz
(b) __________ ° C

10.5 A pipe with one closed end is 13.0 meters long. How many harmonics does it have in the range of human hearing: from 20 Hz to 20,000 Hz? Use 343 m/s for the speed of sound and report the answer to 2 significant figures.

____

10.6 A wire of length 65.0 cm and a mass of 19.6 g is fastened at both ends with tension T. It is plucked, and starts vibrating at its fundamental frequency. A tube 2.15 m long with one closed end has its open end placed right next to the vibrating wire, and by resonance, the air column begins oscillating at its fundamental frequency, which is the same as that of the string. (a) What is the frequency of the vibrating air column? Take the speed of sound to be 343 m/s. (b) Calculate the magnitude of the tension T in the wire.

(a) __________ Hz
(b) __________ N

10.7 A cylindrical drinking glass, 9.40 cm high, is partially filled with liquid. Let x be the depth of the liquid, in cm. A tuning fork vibrating at 1.76 kHz is placed near the open top of the glass. What is the maximum value of x at which the air column above the liquid still resonates with the tuning fork? Take the speed of sound in air to be 343 m/s.

__________ cm

Section 13 - Interactive problem: waves traveling in the same direction

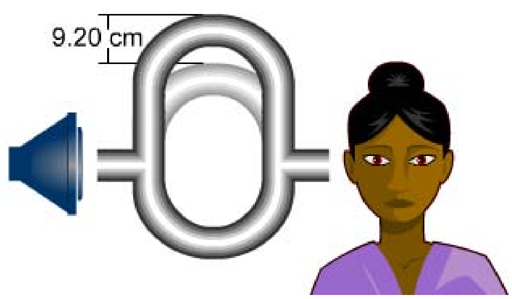
13.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) What phase difference(s) will create a combined wave with the greatest amplitude? (b) What phase difference(s) will create a combined wave that is flat?

(a) ☐ -2π
☐ $-3\pi/2$
☐ $-\pi$
☐ $-\pi/2$
☐ 0
☐ $\pi/2$
☐ π
☐ $3\pi/2$
☐ 2π

(b) ☐ -2π
☐ $-3\pi/2$
☐ $-\pi$
☐ $-\pi/2$
☐ 0
☐ $\pi/2$
☐ π
☐ $3\pi/2$
☐ 2π

Section 14 - Wave interference and path length

14.1 The apparatus shown demonstrates interference of sound waves. A pure tone from the speaker takes both paths to the listener's ear, and the resulting difference in path lengths then creates interference. The length of one path can be changed by sliding the adjustable U-shaped tube at the top. You adjust this tube to find a position that results in complete destructive interference. From this point you then measure how far you must pull the end of the tube up to first hear destructive interference again. If this distance is 9.20 cm, what is the frequency of the sound that is being used? (Use 343 m/s for the speed of sound.)



__________ Hz

14.2 Speaker 1 is positioned at the origin and speaker 2 is at the position (0, 4.00) meters. They emit identical sound waves of wavelength 1.55 m, in phase. If you stand on the *x* axis at (*x*, 0) meters, what is the smallest positive value for *x* for which you experience complete destructive interference?

__________ m

14.3 Two speakers emit sound of identical frequency and amplitude, in phase with each other. The frequency is 415 Hz. If the speakers are 9.00 meters apart, and you stand on the line directly between them, 3.47 m from one speaker, are you standing closer to a point of constructive or destructive interference? (Use 343 m/s for the speed of sound.)

○ Constructive
○ Destructive

14.4 Two speakers are mounted 5.00 meters apart on the ceiling, at a height of 3.05 meters above your ears. You plug up one ear with a cotton ball, and stand with your other ear directly under the midpoint of a line connecting the speakers. The speakers emit identical sound waves, which are in phase when they reach your ear. You start walking, remaining directly under the connecting line, until you first encounter a point of complete destructive interference. You have walked 0.250 meters. What is the frequency of the sound coming from the speakers? (Use 343 m/s for the speed of sound.)

__________ Hz

Section 17 - Beats

17.1 Two violinists are playing their "A" strings. Each is perfectly tuned at 440 Hz and under 245 N of tension. If one violinist turns her peg to tighten her A string to 251 N of tension, what beat frequency will result? Express your answer to the nearest 100^{th} of a Hz.

__________ Hz

17.2 Two identical strings are sounding the same fundamental tone of frequency 156 Hz. Each string is under 233 N of tension. The peg holding one string suddenly slips, reducing its tension slightly, and the two tones now create a beat frequency of three beats per second. What is the new tension in the string that slipped?

__________ N

17.3 Two cellists play their C strings at their fundamental frequency of 65.4 Hz. They are identical strings at the same tension. One of the cellists plays a glissando (slides her finger down the string) until she has shortened it to an effective length of 15/16 the length of the other cellist's C string. What beat frequency will result? Express your answer to the nearest tenth of a Hz.

__________ Hz

17.4 Consider the musical note "A above middle C", known as "concert pitch" or "A440." The frequency of this note is 440 Hz by international agreement. In the chromatic scale, the frequency of a sharp is a factor of 25/24 higher than the note, and the frequency of a flat is a factor of 24/25 lower. If the "A sharp" and the "A flat" notes corresponding to A440 are played together, what will be the resulting beat frequency? (State your answer to the nearest Hz.)

__________ Hz

17.5 You hold two tuning forks oscillating at 294 Hz. You give one of the forks to your friend who walks away at 1.50 m/s. What beat frequency do you hear? Use 343 m/s as the speed of sound and give your answer to the nearest 100^{th} of a Hz.

__________ Hz

17.6 An opera singer walks towards a wall while singing a steady note of 960 Hz. The sound is reflected from the wall back to the opera singer, who hears 10 beats per second in the combination of the sound she sings and the reflected sound. Determine the speed of the walking opera singer.

__________ m/s

17.7 Two pipes − one with two open ends, the other with only one open end − are sounded at the same time. The beat frequency between their fundamental frequencies is 3.00 Hz, and the pipe with two open ends has the higher frequency. The pipe with two open ends is 1.30 meters long. What is the length of the other pipe? Use 343 m/s for the speed of sound.

__________ m

17.8 A small boat approaches a large stationary ship at 4.00 m/s. The boat blows its horn at a frequency of 196 Hz. What is the beat frequency between the ship's horn and its echo against the hull of the ship? State your answer to the nearest 100^{th} of a Hz.

__________ Hz

Chapter 19 Temperature and Heat

19.0 - Introduction

Thermodynamics is the study of heat ("thermo") and the movement of that heat ("dynamics") between objects. A kitchen provides an informal laboratory for the study of thermodynamics. Manufacturers offer numerous kitchen devices designed to facilitate the flow of heat: stovetops and ovens, convection ovens, toasters, refrigerators, and more. Heat flow changes the temperature of what is being cooked or cooled, and that can be monitored with thermometers.

This chapter starts with a few basic thermodynamics concepts, namely how temperature is measured, what temperature scales are, and what is meant by heat. It then begins the discussion of the relationship between heat and temperature.

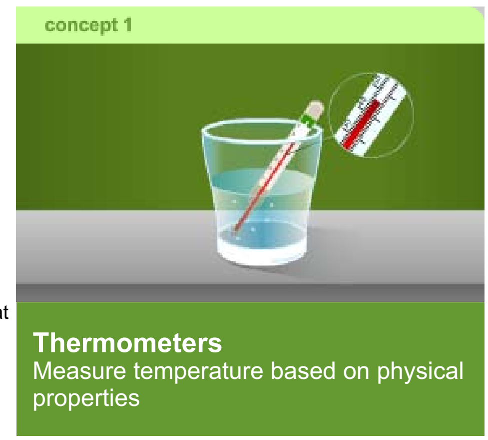
Kitchen appliances are engineered using the principles of thermodynamics.

19.1 - Temperature and thermometers

Although temperature is an everyday word, like energy it is surprisingly hard to define. For now, we ask that you continue to think of temperature simply as something measured by a *thermometer*. Warmer objects have higher temperatures than cooler objects.

Traditional thermometers rely on the important principle that any two objects placed in contact with each other will reach a common temperature. For instance, when a traditional fever thermometer is placed under your tongue, after a few minutes the flow of heat causes it to reach the same temperature as your body.

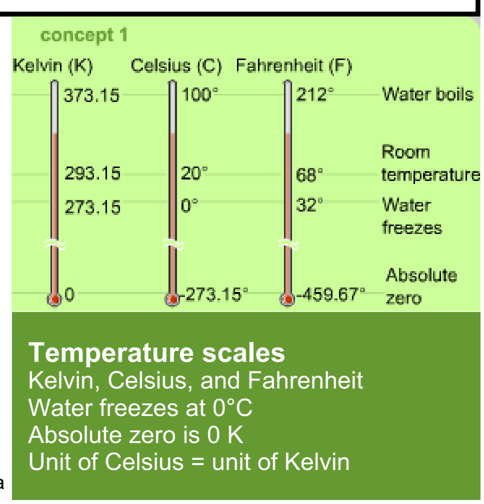
While there are many different types of thermometers available, they all rely on some physical property of materials in order to measure temperature. In the "old" days, body temperature was measured with a glass thermometer filled with mercury, a material that expands significantly with temperature and whose expansion is proportional to the change in temperature.



Today, a wide variety of physical properties are used to determine temperature. Some medical clinics use thermometers that measure temperature with plastic sheets containing a chemical that changes color with temperature. Battery-powered digital thermometers rely on the fact that a resistor's resistance changes with temperature. Ear thermometers use *thermopiles* that are sensitive to subtle changes in the infrared radiation emitted by your body; this radiation changes with your temperature.

19.2 - Temperature scales

In the United States, the Fahrenheit system is the most common measurement system for temperature. The units in this system are called degrees. In most of the rest of the world, however, temperatures are measured in degrees Celsius. Physicists use the Celsius scale or, quite often, another scale called the Kelvin scale. All three scales are shown on the right.



There are two things required to construct a temperature scale. One is a reference point, such as the temperature at which water freezes at standard atmospheric pressure. As shown to the right, the three scales have different values at this reference point. Water freezes at 273.15 kelvins (273.15 K), $0°$ Celsius $(0°C)$ and $32°$ Fahrenheit $(32°F)$. Notice that the standard terminology for the Kelvin scale avoids the use of "degrees." Water freezes at 273.15 kelvins, **not** 273.15 degrees Kelvin.

The other requirement is to pick another reference point, such as the temperature at which water boils at standard atmospheric pressure, and establish the number of degrees between these two points. This determines the magnitude of the units of the scale. The Celsius and Kelvin scales both have 100 units between the freezing and boiling points of water. This means that their units are equal: a change of $1\ C°$ equals a change of 1 K. (Changes in Celsius temperatures are indicated with $C°$ instead of $°C$.) In contrast, there are 180 degrees between these temperatures in the Fahrenheit system.

Another important concept is shown in the illustration to the right: absolute zero. At this temperature, molecules (in essence) cease moving. Reaching this temperature is not theoretically possible, but temperatures quite close to this are being achieved. Absolute zero is 0 K, or $-273.15°C$.

To standardize temperatures, scientists have agreed on a common reference point called the *triple point*. The triple point is the sole combination of pressure and temperature at which solid water (ice), liquid water, and gaseous water (water vapor) can coexist. It equals 273.16 K at a pressure of 611.73 Pa. The triple point is used to define the kelvin as an SI unit. One kelvin equals 1/273.16 of the difference between absolute zero and the triple point.

If you are a sharp-eyed reader, you may have noticed the references to both 273.16 and 273.15 in this section. The freezing point of water is typically stated as 273.15 K $(0°C)$ because this is its value at standard atmospheric pressure, but at the triple point pressure, water freezes at 273.16 K $(0.01°C)$.

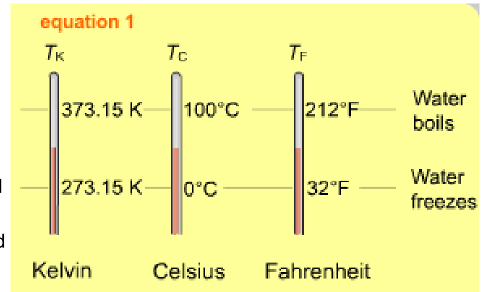
19.3 - Temperature scale conversions

Since the Celsius and Kelvin scales have the same number of units between the freezing and boiling points of water, it takes just one step to convert between the two systems, as you see in the first conversion formula in Equation 1. To convert from degrees Celsius to kelvins, add 273.15. To convert from kelvins to degrees Celsius, subtract 273.15.

Since water freezes at $32°$ and boils at $212°$ in the Fahrenheit system, there are 180 degrees Fahrenheit between these points, compared to the 100 units in the Celsius and Kelvin systems. To convert from degrees Fahrenheit to degrees Celsius, first subtract 32 degrees (to establish how far the temperature is from the freezing point of water) and then multiply by 100/180, or 5/9, the ratio of the number of degrees between freezing and boiling on the two systems. That conversion is shown as the second equation in Equation 1. If you further needed to convert to kelvins, you would add 273.15.

To switch from Celsius to Fahrenheit, you first multiply the number of degrees Celsius by 9/5 (the reciprocal of the ratio mentioned above) and then add 32.

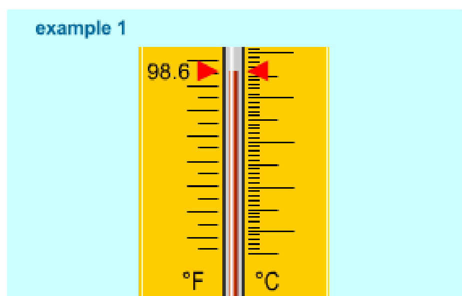
In Example 1, you see the conventionally normal human body temperature, $98.6°F$, converted to degrees Celsius and kelvins.



Temperature scales: conversions

$$T_K = T_C + 273.15$$
$$T_C = (5/9)(T_F - 32)$$

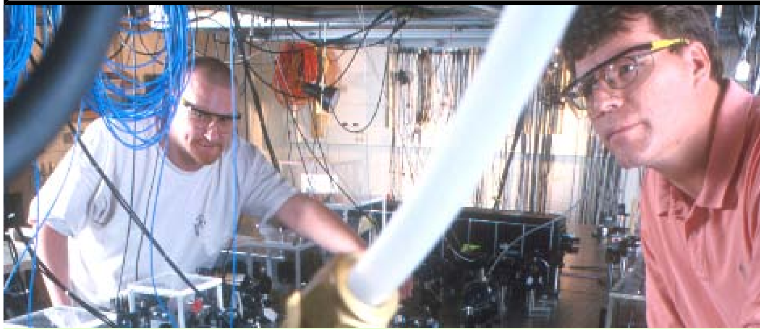
T_K = Kelvin temperature
T_C = Celsius temperature
T_F = Fahrenheit temperature



Convert 98.6°F to Celsius and Kelvin.

$T_C = (5/9)(T_F - 32.0)$
$T_C = (5/9)(98.6°F - 32.0)$
$T_C = (5/9)(66.6) = 37.0°C$
$T_K = T_C + 273.15$
$T_K = 37.0 + 273.15 = 310\text{ K}$

19.4 - Absolute zero



concept 1

Absolute zero
It cannot get colder
Molecules at minimum energy state
0 K, –273.15°C

Absolute zero: As cold as it can get.

Absolute zero is a reference point at which molecules are in their minimum energy state (quantum theory dictates they still have some energy). It does not get colder than absolute zero; nothing with a temperature less than this minimum energy state can exist.

Physics theory says it is impossible for a material to be chilled to absolute zero. Instead, it is a limit that scientists strive to get closer and closer to achieving. Above, you see a photograph of scientists who chilled atoms to less than a hundred-billionth of a degree above absolute zero. At that temperature, the atoms changed into a state of matter called a Bose-Einstein condensate.

Although we exist around the relatively toasty 293 K thanks to the Sun and our atmosphere, the background temperature of the universe − the temperature far from any star − is only about 3 K. Brrrr.

19.5 - Heat

Heat: Thermal energy transferred between objects because of a difference in their temperatures.

If you hold a cold can of soda in your hand, the soda will warm up and your hand will chill. Energy flows from the object with the higher temperature − your hand − to the object with the lower temperature − the soda. This energy that moves is called heat. Ovens are designed to facilitate the flow of heat. In the diagram to the right, you see heat flowing from the oven coils through the air to the loaf of bread. As with other forms of energy, heat can be measured in joules. It is represented by the capital letter Q.

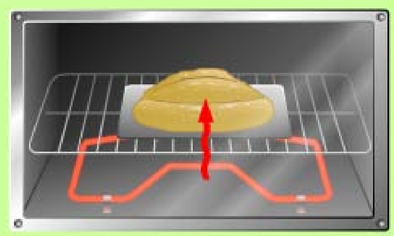
Physicists do **not** say an object has heat. Heat refers solely to the flow of energy due to temperature differences. Heat transfers *thermal energy* that is internal to objects, related to the random motion of the atoms making up the objects.

Heat is like work: It changes the energy of an object or system. It does not make sense to say “how much work a system has”, nor does it make sense to say “how much heat the system has”. Just as work is done by a system or on a system, heat as thermal energy can enter a system or leave a system.

Having said that heat is measured in joules, we will backtrack a little in order to explain some other commonly used units. These units measure heat by its ability to raise the temperature of water. For example, a *calorie* raises one gram of water from $14.5°C$ to $15.5°C$. The *British Thermal Unit* (BTU) measures the heat that would raise a pound of water $1°$ on the Fahrenheit scale.

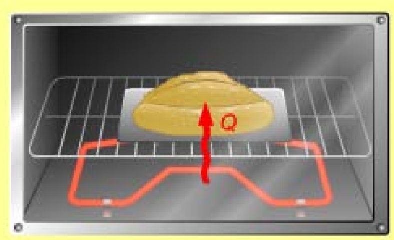
The unit we use to measure one property of food − the calories you see labeled on the back of food packages − is actually a kilocalorie (good marketing!). This is sometimes spelled with a capital C, as in Calories. Food calories measure how much heat will be released when an object is burned. A Big Mac® hamburger contains 590 Calories, or 590,000 calories. This amount of energy equals about 2,500,000 J. If your body could capture all this energy, if it were 100% efficient and solely focused on the task, the energy from a Big Mac would be enough to allow you to lift a 50 kg weight one meter about 5000 times. We will skip the calculation for the french fries.

concept 1

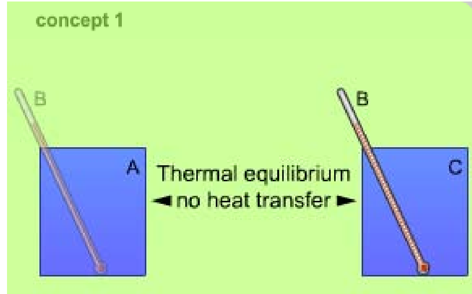


Heat
Energy flow due to temperature difference
Not a property of an object

equation 1



Heat
Q represents heat
Units: joules (J)



Zeroth law of thermodynamics
If A, B in thermal equilibrium,
and B, C in thermal equilibrium,
then A, C in thermal equilibrium
(no heat transfer)

Zeroth law of thermodynamics: If objects A and B are in thermal equilibrium, and objects B and C are in thermal equilibrium, then A and C will be in equilibrium as well.

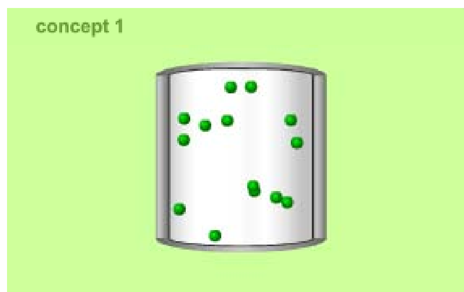
When you place two objects with different temperatures next to each other, the warmer object will cool off and the cooler object will warm up. Heat will flow until the objects reach *thermal equilibrium*, meaning they have the same temperature. For instance, place a pint of ice cream in a warm car, and the result will be warmer ice cream and a cooler car.

Thermometers rely on heat flowing until they reach thermal equilibrium with the substance whose temperature they are measuring. Their practical use also relies on another principle, called the zeroth law of thermodynamics. This principle states that if object A is in thermal equilibrium with object B, and object B is in equilibrium with object C, then A and C will be in equilibrium when they are placed in direct contact, and no heat will flow between them.

We illustrate this law on the right. Let's say you put thermometer B in a container of water A. When the thermometer's reading stabilizes at a constant value, say $20°\text{C}$, it has reached thermal equilibrium with the water. If you then place the thermometer in a second container C and its reading remains $20°\text{C}$ there, you can conclude that A and C would be in thermal equilibrium when placed in direct contact with each other. They have the same temperature and heat would not flow between them.

This may seem commonsense, but it is an important assumption in thermodynamics. Its importance was realized after the first and second laws of thermodynamics (which you will study later) were already codified − hence it became the zeroth law, since it is an underlying assumption for the other laws.

19.7 - Internal energy



Internal energy
Energy of system's atoms, molecules

Internal energy: The energy associated with the molecules and atoms that make up a system.

In the study of mechanics, energy is an overall property of an object or system. The energy is a function of factors like how fast a car is moving, how high an object is off the ground, how fast a wheel is rotating, and so forth.

In thermodynamics, the properties of the molecules and/or atoms that make up the object or system are now the focus. They also have energy, a form of energy called internal energy. The internal energy includes the rotational, translational and vibrational energy of individual molecules and atoms. It also includes the potential energy within and between molecules.

To contrast the two forms of energy: If you lift a pot up from a stovetop, you will increase its gravitational potential energy. But in terms of internal energy, nothing has changed. The potential energy of the pot's molecules based on their relationship to each other has not changed.

However, if you turn on the burner under the cooking pot, the flow of heat will increase the kinetic energy of its molecules. The molecules will move faster as heat flows to the pot, which means the internal energy of the molecules of the pot increases.

19.8 - Thermal expansion

Expansion joints allow bridge sections to expand without breaking.

Thermal expansion: The increase in the length or volume of a material due to a change in its temperature.

You buy a jar of jelly at the grocery store and store it on a pantry shelf. When it comes time to open the jar, the lid refuses to budge. Fortunately, you know that placing the jar under hot water will increase your odds of being able to twist open the lid.

By using hot water to coax a lid to turn, you are implicitly using two physics principles. First, most materials expand as their temperature increases. Second, different materials expand more or less for a given increase in temperature. The metal lid of the jar expands more than the glass container as you increase their temperatures, effectively lessening the "grip" of the lid on the glass. (The temperature of the lid is also likely to increase faster, another factor that accounts for the success of the process.)

The expansion of materials due to a temperature change can be useful at times, as the jar-opening example demonstrates. Sometimes, it poses challenging engineering problems. For example, when nuclear waste is stored in a rock mass, heat can flow from the waste to the rock, raising the rock's temperature and causing it to expand and crack. This could allow the dangerous waste to leak out. Knowing the exact rate of expansion can help engineers design storage intended to prevent cracking.

Good engineering takes expansion into account. For instance, bridges are built with expansion joints, like the one shown at the top of this page, that accommodate expansion as the temperature increases.

Thermal expansion
Most materials expand with increased temperature
Different materials expand at different rates

19.9 - Thermal expansion: linear

Thermal linear expansion: Change in the length of a material due to a change in temperature.

Coefficient of linear expansion (1/C°)	
Carbon steel	1.17×10^{-5}
Iron	1.18×10^{-5}
Copper	1.65×10^{-5}
Silver	1.89×10^{-5}
Aluminum	2.31×10^{-5}
Magnesium	2.48×10^{-5}
Lead	2.89×10^{-5}
	at temperatures near 25°C

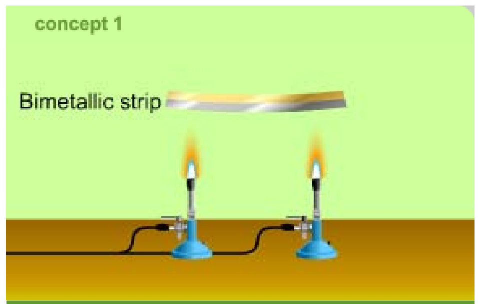
Most objects expand with increased temperature; how much they expand varies by material. In this section, we discuss how much they expand in one dimension, along a line. Their expansion is measured as a fraction of their initial length.

In Equation 1, you see the equation for linear expansion. The change in length equals the initial length, times a constant α (Greek letter alpha), times the change in temperature. The constant α is called the *coefficient of linear expansion* and depends on the material. A table of coefficients of linear expansion for some materials is shown above. These coefficients are valid for temperatures around $25°\mathrm{C}$.

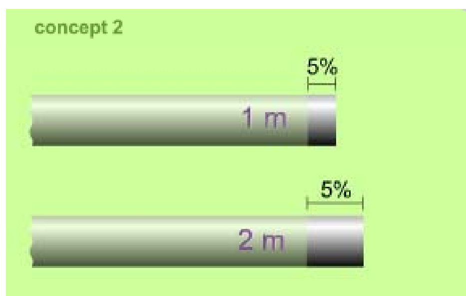
Differing coefficients of linear expansion can be taken advantage of to build useful mechanisms. A *bimetallic strip*, shown to the right, consists of two metals with different coefficients of linear expansion. As the strip increases in temperature, the two materials expand at different rates, causing the strip to bend. Since the amount of bending is a function of the temperature, such a strip can be used in a thermometer to indicate temperature. It can also be used as a thermostat to control appliances, such as coffee pots and toasters. In these appliances, the bending of the strip interrupts a circuit and turns off the power when the appliance has reached a specified temperature.

Significant changes in temperature cause fairly minor changes in length. For instance, in Example 1, we calculate the expansion of a 0.50 meter copper rod when its temperature is increased $80\ \mathrm{C}°$. The increase in length is just 6.6×10^{-4} meters, less than a millimeter.

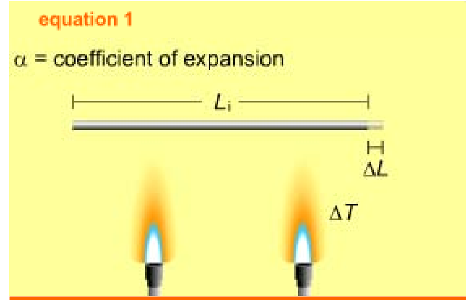
Some materials, like carbon fiber, have negative coefficients of expansion, which means they shrink when their temperature increases. By blending materials with both positive and negative coefficients, engineers design systems that change shape very little with changes in temperature. The Boeing Company pioneered the use of negative coefficient materials in airplanes and satellites.



Thermal expansion: linear
Measured along one dimension
Constant α depends on material



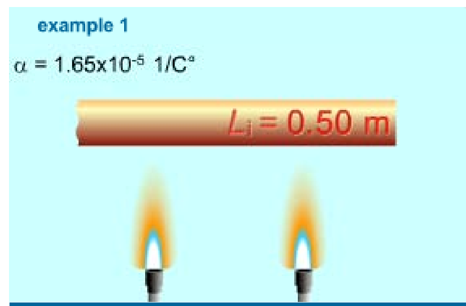
Rods of same material
Expansion proportional to initial length



Linear expansion

$$\Delta L = L_i \alpha \Delta T$$

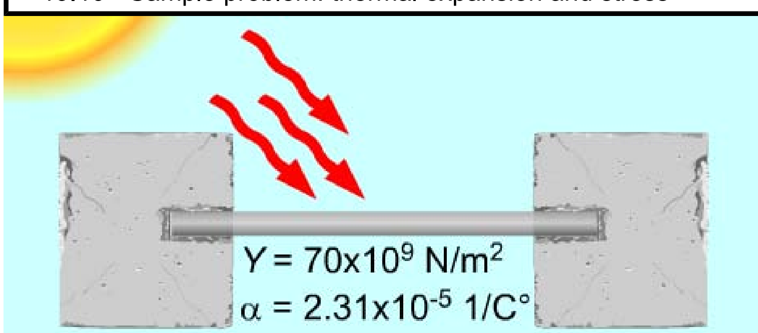
L = length
α = coefficient of linear expansion
ΔT = change in temperature
Coefficient calibrated for K or °C



The copper rod is heated from 15°C to 95°C. What will its increase in length be?

$\Delta L = L_i \alpha \Delta T$
$\Delta T = 95°C - 15°C = 80$ C°
$\Delta L = (0.5 \text{ m})(1.65\times10^{-5} \text{ 1/C°})(80 \text{ C°})$
$\Delta L = 6.6\times10^{-4}$ m

19.10 - Sample problem: thermal expansion and stress



What stress does the aluminum rod exert when its temperature rises 20 K?

Above, you see an aluminum rod heated by the Sun and held in place with concrete blocks. Since the rod increases in temperature, its length also increases. This exerts a force on the concrete blocks. Stress is force per unit area, and an equation for tensile stress was presented in another chapter. Young's modulus for aluminum is given; it relates the fractional increase in length (the strain) to stress. You are asked to find the stress that results from the increase in temperature.

Variables

thermal expansion coefficient	$\alpha = 2.31\times10^{-5}$ 1/C°
Young's modulus	$Y = 70\times10^{9}$ N/m^2
temperature change	$\Delta T = 20$ K
initial length	L_i
change in length	ΔL
tensile stress	F/A

You may notice that the initial length of the rod is not known. It is not needed to answer the question.

What is the strategy?

1. Combine the equations for thermal expansion and tensile stress to write an equation to calculate tensile stress from the temperature change.
2. Use the equation to compute the stress in this case.

Physics principles and equations

We will use the equations for thermal expansion and tensile stress. Tensile stress is measured as force per unit area, or F/A.

$$\Delta L = L_i \alpha \Delta T$$

$$F/A = Y(\Delta L/L_i)$$

Step-by-step solution

We start by substituting the expression for the change in length from the thermal expansion equation into the tensile stress equation, and then do some algebraic simplification.

Step	Reason
1. $F/A = Y(\Delta L/L_i)$	tensile stress equation
2. $\Delta L = L_i \alpha \Delta T$	thermal expansion equation
3. $F/A = Y\dfrac{L_i \alpha \Delta T}{L_i}$	Substitute equation 2 into equation 1
4. $F/A = Y\alpha\Delta T$	L_i cancels out

The equation we just found does not depend on the initial length. We know all the values needed to calculate the tensile stress in the aluminum rod.

Step	Reason
5. $F/A = (70\times10^{9}\text{ N/m}^2)(2.31\times10^{-5}\text{ 1/C°})(20\text{ K})$	enter values
6. $F/A = 3.2\times10^{7}$ N/m^2	multiply

An aluminum rod with a radius of 0.025 meters (about one inch) exerts more than 60,000 newtons of force (equivalent to the weight of a large elephant) against perfectly rigid supports when its temperature increases 20 C°! In this problem, we have ignored the expansion of the concrete in which the aluminum is embedded.

19.11 - Thermal expansion of water

Water exhibits particularly interesting expansion properties when it nears its freezing point. Above 4°C, water expands with temperature, as most liquids do. However, water also expands as it **cools** from 4°C to 0°C, a significant and unusual phenomenon. Below 0° water once again contracts as it cools. The consequence is that liquid water is most dense at a temperature of around 4°C.

This pattern of expansion means that lakes and other bodies of water freeze from the top down. Why? In colder climates, as the autumn or winter seasons approach and the air temperature drops to near freezing or below, the water in a lake cools. When the water is cold but still above 4°C, it contracts when it chills. Since it becomes denser, it sinks. This brings warmer water to the surface, which cools in turn. Eventually, the entire lake reaches 4°C. But when the top layer then becomes colder still, it no longer sinks. Below 4°C, the water expands, becoming less dense. It floats atop the warmer water.

As the cold water at the top of the lake further cools and freezes, it forms a floating layer

concept 1

Thermal expansion of water
Water contracts and sinks as it cools,

of ice that insulates the water below. Water is also atypical in that its solid form, ice, is less dense than its liquid form and floats on top of it. Fish and other aquatic life can live in the relatively warm (and liquid) water below, protected by a shield of ice.

If water always expanded with increasing temperature for all temperatures above 0°C, and contracted with decreasing temperature, the coldest water would sink to the bottom where it might never warm up. Water's negative coefficient of expansion in the temperature range from 0°C to 4°C is crucial to life on Earth. If ice did not float, oceans and lakes would freeze from the bottom to the top. This would increase the likelihood that they would freeze entirely, since they would not have a top layer of ice to insulate the liquid water below and their frozen depths would not be exposed to warm air during the spring and summer.

until 4°C
From 4°C to 0°C, water expands and stays on top
Then ice forms on top and floats

19.12 - Thermal expansion: volume

Thermal volume expansion: Change in volume due to a change in temperature.

Coefficient of volume expansion (1/C°)

Liquids		*Solids*	
Mercury	19.6×10^{-5}	Glass*	2.14×10^{-5}
Water	20.7×10^{-5}	Copper*	5.00×10^{-5}
Glycerin	50.4×10^{-5}	Silver*	5.64×10^{-5}
Olive Oil	72.0×10^{-5}	Lead*	8.37×10^{-5}
Methyl Alcohol	120×10^{-5}	Ice (-26°C)	11.3×10^{-5}
Acetone	149×10^{-5}		

* between 0 -100°C

The equation for thermal linear expansion is used to calculate the thermally induced change in the size of an object in just one dimension. Thermal expansion or contraction also changes the volume of a material, and for liquids (and many solids) it is more useful to determine the change in volume rather than expansion along one dimension. The expansion in volume can be significant. Automobile cooling systems have tanks that capture excess coolant when the heated fluid expands so much it exceeds the radiator's capacity. A radiator and its overflow tank are shown in Concept 1 on the right.

The formula in Equation 1 resembles that for linear expansion: The increase is proportional to the initial volume, a constant, and the change in temperature. The constant β is called the *coefficient of volume expansion*.

Above, you see a table of coefficients of volume expansion for some liquids and solids. The coefficients for liquids are valid for temperatures at which these substances remain liquid.

For solid materials like copper and lead, the coefficient of volume expansion β is about three times the coefficient of linear expansion α, because the solid expands linearly in three dimensions.

concept 1

Thermal expansion: volume
Volume increases with temperature
Constant β varies by material
Increase is proportional to initial volume

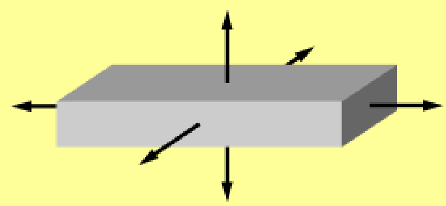
equation 1

Thermal expansion: volume

$$\Delta V = V_i \beta \Delta T$$

V = volume
β = coefficient of volume expansion
ΔT = change in temperature
Coefficient calibrated for K or °C

equation 2



For solids

$\beta \approx 3\alpha$
β = coefficient of volume expansion
α = coefficient of linear expansion

example 1

The temperature of 2.0 L of water increases from 5.0° C to 25° C. How much does its volume increase?

$\Delta V = V_i \beta \Delta T$

ΔT = 25°C – 5.0°C = 20 C°

ΔV = (2.0 L)(2.07×10^{-4} 1/C°)(20 C°)

ΔV = 0.0083 L

19.13 - Interactive checkpoint: a radiator

A truck has a radiator that holds 0.0176 m^3 of coolant. The coefficient of volume expansion of the coolant is the same as that of water, 2.07×10^{-4} (1/C°). The truck starts a trip with a full radiator at 18.0°C. After 30 minutes, the coolant temperature in the radiator is 102.0°C. What volume of coolant has flowed into the radiator's overflow container? (Ignore any expansion of the radiator.)

Answer:

ΔV = ______ m^3

19.14 - Specific heat

Specific heat: A proportionality constant that relates the amount of heat flow per kilogram to a material's change in temperature.

	Specific heat (J/kg·K)
Lead	129
Silver	235
Copper	385
Iron	449
Carbon	709
Aluminum	897
Air (27°C)	1007
Ice (0°C)	2110
Water (30°C)	4178
	at 10^5 Pa, 25°C

Specific heat is a property of a material; it is a proportionality constant that states a relationship between the heat flow per kilogram of a material and its change in temperature.

Heat capacity and specific heat are closely related. Heat capacity is a property of an object, which has a particular size and consists of one or more materials; specific heat is a property of a material. The heat capacity of an object consisting of a single material equals the specific heat of the material times its mass.

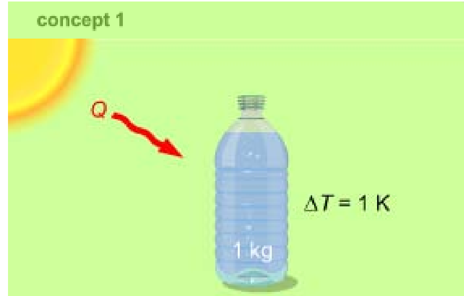
A material's specific heat is determined by how much heat is required to increase the temperature of one kilogram of the material by one kelvin.

A material with a greater specific heat requires more heat per kilogram to increase its temperature a given amount than one with a lesser specific heat. In spite of its name, specific heat is not an amount of heat, but a constant relating heat, mass, and temperature change.

The specific heat of a material is often used in the equation shown in Equation 1. The heat flow equals the product of a material's specific heat c, the mass of an object consisting of that material, and its change in temperature. The illustration in Equation 1 shows how specific heat relates heat flow to change in temperature. As you can see from the graph, lead increases in temperature quite readily when heat flows into it, because of its low specific heat.

In contrast, water, with a high specific heat, can absorb a lot of energy without changing much in temperature. Temperatures in locations at the seaside, or having humid atmospheres, tend to change very slowly because it takes a lot of heat flow into or out of the water to accomplish a small change in temperature. Summer in the desert southwest of the United States is famous for its blazing hot days and chilly nights, while on the east coast of the country the sweltering heat of the day persists long into the night. Materials with large specific heats are sometimes informally called "heat sinks" because of their ability to store large amounts of internal energy without much temperature change.

Above, you see a table of some specific heats, measured in joules per kilogram·kelvin. The specific heat of a material varies as its temperature and pressure change. The table lists specific heats for materials at $25°C$ to $30°C$ (except for ice) and 10^5 Pa pressure, about one atmosphere. Specific heats vary somewhat with temperature, but you can use these values over a range of temperatures you might encounter in a physics lab (or a kitchen).

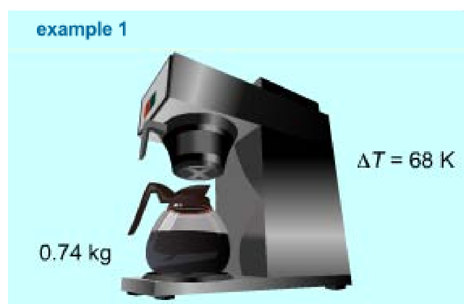


Specific heat of a material
Relates heat and temperature change, per kilogram

Specific heat of a material

$$Q = cm\Delta T$$

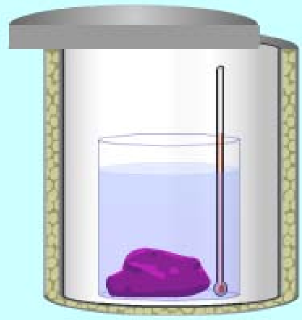
Q = heat
c = specific heat (J/kg·K)
m = mass
ΔT = temperature change in C° or K



How much heat is required to increase the coffee's temperature 68 K?

$Q = cm\Delta T$
Q = (4178 J/kg·K)(0.74 kg)(68 K)
Q = 210,000 J

19.15 - Sample problem: a calorimeter



A calorimeter is used to measure the specific heat of an object. The water bath has an initial temperature of 23.2°C. An object with a temperature of 67.8°C is placed in the beaker. After thermal equilibrium is reestablished, the water bath's temperature is 25.6°C. What is the specific heat of the object?

In a calorimeter, a water bath is placed in a well-insulated container. The temperature of the water bath is recorded, and an object of known mass and temperature placed in it. After thermal equilibrium is reestablished, the temperature is measured again. From this information, the specific heat of the object can be calculated. (We ignore the air in this calculation.)

The use of a calorimeter depends on the conservation of energy. In the calorimeter, heat flows from the object to the water bath (or vice-versa if the object is colder than the water). Because the calorimeter is well insulated, negligible heat flows in or out of it. The conservation of energy allows us to say that the heat lost by the object equals the heat gained by the water bath.

The water bath consists of the water and the beaker containing it. If the mass of the water is much greater that the mass of the beaker, relatively little heat will be transferred to the beaker and it can be ignored in the calculation. We do that here.

Variables

	water	object
mass	$m_w = 0.744$ kg	$m_o = 0.197$ kg
initial temperature	$T_w = 23.2$°C	$T_o = 67.8$°C
specific heat	$c_w = 4178$ J/kg·K	c_o
heat transferred to water	Q_w	
heat transferred from object	Q_o	
final temperature	$T_f = 25.6$°C	

What is the strategy?

1. Use conservation of energy to state that the heat lost by the object equals the heat gained by the water bath. Apply the specific heat equation to write the conservation of energy equation in terms of the masses, temperatures, and specific heats.
2. Solve for the unknown specific heat, substitute the known values, and evaluate.

Physics principles and equations

By the conservation of energy, the heat gained by the water bath (beaker plus water) equals the heat lost by the object. The sum of the heat transfers is zero.

$$Q_w + Q_o = 0$$

We use the equation below to relate the heat flow to specific heat.

$$Q = cm\Delta T$$

Step-by-step solution

We start with an equation stating the conservation of energy for this experiment. Then, the heat values in the equation are written with expressions involving specific heat.

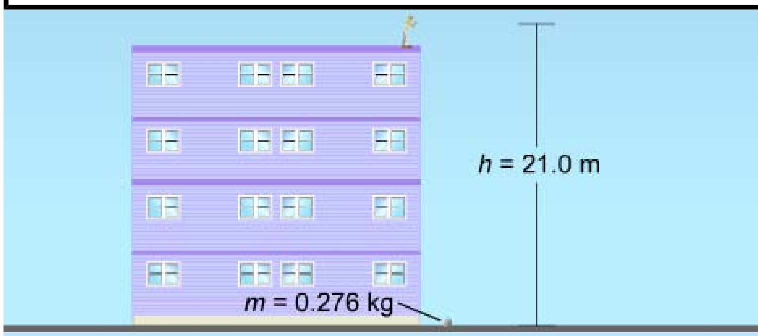
Step	Reason
1. $Q_w + Q_o = 0$ $Q_w = -Q_o$	conservation of energy
2. $Q_o = c_o m_o \Delta T$	specific heat equation
3. $Q_o = c_o m_o (T_f - T_o)$	initial and final temperatures for object
4. $Q_w = c_w m_w (T_f - T_w)$	specific heat equation
5. $c_w m_w (T_f - T_w) = -c_o m_o (T_f - T_o)$	substitute equations 3 and 4 into equation 1

Now we solve the equation for the unknown specific heat of the object and evaluate.

Step	Reason
6. $c_o = -\dfrac{c_w m_w (T_f - T_w)}{m_o (T_f - T_o)}$	solve for specific heat of object
7. $c_o = -\dfrac{(4178\ \text{J/kg}\cdot\text{K})(0.744\ \text{kg})(25.6\ °\text{C} - 23.2\ °\text{C})}{(0.197\ \text{kg})(25.6\ °\text{C} - 67.8\ °\text{C})}$	substitute
8. $c_o = 897$ J/kg·K	evaluate

Based on the values in the table of specific heats, it appears that the material may consist of aluminum.

19.16 - Interactive checkpoint: KE to heat



A 0.276 kg lead ball falls from a height of 21.0 m and lands on the ground without bouncing. Assume that half the energy generated by the impact of the ball with the ground becomes internal energy in the ball. The specific heat of lead is 129 J/kg·K. What is the temperature change of the ball?

Answer:

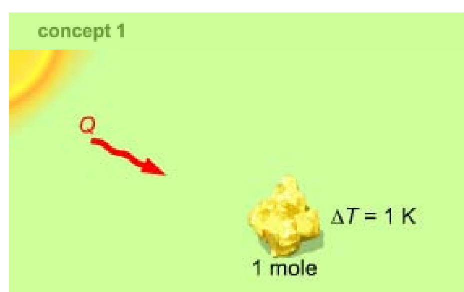
$\Delta T =$ ______ K

19.17 - Molar specific heat

Molar specific heat: A proportionality constant that relates the amount of heat flow per mole to a material's change in temperature.

Molar specific heat	(J/mol·K)
Aluminum	24.20
Copper	24.44
Iron	25.10
Silver	25.35
Lead	26.65
	at 10^5 Pa, 25°C

Scientists find it convenient at times to measure substances in terms of moles. If you have studied chemistry, you probably studied moles. Briefly, one mole of a substance contains 6.022×10^{23} particles (typically molecules; moles are explained in more depth later). Measuring in moles focuses particularly on the number of molecules in an object instead of its mass.



Molar specific heat
Relates heat and temperature change, per mole

Specific heat and molar specific heat are both proportionality constants, relating the heat transfer per an amount of a material to the resulting change in temperature. Specific heat is stated in terms of joules per kilogram, and molar specific heat in terms of joules per mole.

A material's molar specific heat is determined by how many joules are required to heat one mole of the substance one kelvin. A material with a greater molar specific heat requires more heat per mole to produce a given change in temperature than a material with a lesser molar specific heat. This is quantified in Equation 1.

The table above lists the molar specific heats of some metals at room temperature. Measuring specific heat in terms of moles reveals an interesting fact: The values do not vary much. In fact, the molar specific heats of all solids approach a value of about 25 J/mol·K as their temperatures increase. When they turn into liquids or gases, their molar specific heats change.

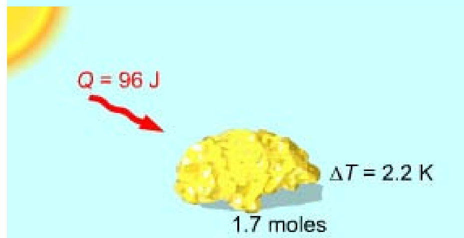
This consistency means that the differences in specific heat values for solids (when measuring by kilograms) are due mainly to the number of molecules contained in a kilogram, rather than differences in the properties of the solids.

equation 1

$$Q = kn\Delta T$$

Q = heat
k = molar specific heat (J/mol·K)
n = number of moles
ΔT = temperature change in C° or K

example 1



This is a gold nugget. Its temperature increases 2.2 K when 96 J of heat are added. What is the molar specific heat of gold?

$Q = kn\Delta T$

$k = Q/n\Delta T$

$k = (96\text{ J})/(1.7\text{ mol})(2.2\text{ K})$

$k = 26\text{ J/mol·K}$

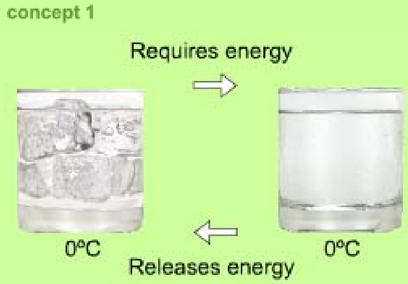
19.18 - Phase changes

Phase change: Transformation between solid and liquid, liquid and gas, or solid and gas.

When you pop some ice cubes into a drink, they will melt. Heat flows from the warmer drink to the cooler ice cubes. Let's say the ice cubes start at $-10°\text{C}$, cooler than the freezing point of water, and they are dropped into a pot of hot coffee. Initially, heat flowing to the ice cubes raises their temperature. But at $0°\text{C}$, heat will flow to the ice cubes from the still warm coffee without the cubes changing temperature. That is, it takes energy to liberate the water molecules from the crystal structure of the ice and allow them to move freely at the same temperature through the coffee. This occurs as the ice melts, changing phase from a solid to a liquid. Phase changes between solid, liquid and gas do not change an object's temperature, but they do require heat transfer.

Phase changes occur as heat flows into or out of a substance. An ice cube melts in hot coffee, but the icemaker in a freezer causes water to change from a liquid to ice. In a freezer, heat is transferred from the liquid water to the cooler freezer.

concept 1



Phase change
Transformation between states
Consumes energy or releases energy
Temperature stays constant

In days of yore, refrigerators were called "iceboxes" because ice was used to cool the contents of the box. Heat would flow from the warmer air to the cooler ice, cooling the air. As the ice warmed and then melted, or changed phase, it would be replaced with a new block. Modern refrigerators continue to use phase changes (between liquid and gas), but they employ substances other than water.

The temperature at which a substance changes phase depends on the substance. For instance, water melts at $0°\text{C}$ at atmospheric pressure, but iron melts at $1538°\text{C}$.

Some substances can "skip" the liquid state by transforming directly from a solid to a gas or vice-versa. This is called *sublimation*. Mothballs sublimate: They transform from a solid directly into a gas. "Dry" ice (solid carbon dioxide) is another solid that sublimates directly into gas at atmospheric pressure. Frost in your freezer is an example of sublimation in the reverse direction. In this case, gaseous water vapor changes directly into solid ice.

Latent heat: Energy required per kilogram to cause a phase change in a given material.

	Melting point (°C)	Latent heat of fusion (J/kg)	Boiling point (°C)	Latent heat of vaporization (J/kg)
Aluminum	660	3.97×10^5	2519	1.09×10^7
Carbon	4489	9.74×10^6		
Copper	1085	2.09×10^5		
Iron	1538	2.47×10^5		
Lead	327	2.30×10^4	1749	8.66×10^5
Mercury	-39	1.14×10^4	357	2.95×10^5
Nitrogen	-210	2.53×10^4	-196	1.99×10^5
Table salt	801	3.78×10^5		
Water	0	3.34×10^5	100	2.26×10^6
			at standard pressure	

Latent heats of fusion and vaporization.

Heat flow can cause a substance to change phases by converting it between a solid and a liquid, or a liquid and a gas.

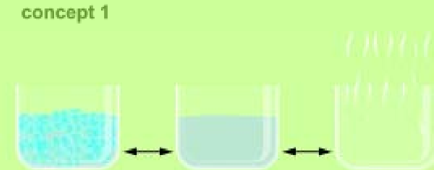
Latent heat describes how much energy per kilogram is required for a given substance to change phase. It is a proportionality constant, expressing the relationship between heat and mass as shown in Equation 1. The constant depends on the material and on the phase change. Different amounts of energy are required to transform a material between its liquid and solid states than between its liquid and gaseous states.

The *latent heat of vaporization* is the amount of heat per kilogram consumed when a given substance transforms from a liquid into a gas, or released when the substance transforms from a gas back to a liquid. The *latent heat of fusion* is the heat flow per kilogram during a change in phase between a solid and a liquid.

The table above shows the latent heats of fusion and vaporization for various substances. For instance, you need 3.34×10^5 J of energy to convert a kilogram of ice (at $0°C$) to liquid water. Continued flow of heat into the water will raise its temperature until it reaches $100°C$. At this temperature, it will take 2.26×10^6 joules of heat to turn it into a gas, about seven times as much as it took to convert it to a liquid.

Salt causes ice to melt, a phenomenon called "freezing point depression." When you add rock salt to the crushed ice in a hand-cranked ice cream freezer, you force the ice to melt. Heat flows from the resulting saltwater solution into the ice as it changes phase from solid to liquid, resulting in a slurry having a temperature far colder than $0°C$. Heat then flows from the ice cream solution into this mixture, and the ice cream freezes.

concept 1



Latent heat
Energy required per kg to change state
Latent heat of fusion: solid to liquid
Latent heat of vaporization: liquid to gas
Amount same in either "direction"

equation 1

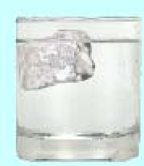
Heat required for phase changes

$$Q = L_f m$$
$$Q = L_v m$$

Q = heat
m = mass
L_f = latent heat of fusion (J/kg)
L_v = latent heat of vaporization (J/kg)

example 1

$L_f = 3.34 \times 10^5$ $\quad c = 4178$ J/kg·K



A glass contains 0.0370 kg of ice at 0°C. How much heat transfers to the ice as it melts without changing temperature?

$Q = L_f m$
$Q = (3.34 \times 10^5 \text{ J/kg})(0.0370 \text{ kg})$
$Q = 1.23 \times 10^4$ J

The glass contains 0.160 kg of water at 30.0°C and 0.0370 kg of ice at 0.00°C. What is the resulting temperature of the water at thermal equilibrium after the ice melts?

In the insulated container, the only source of the heat to melt the ice is the surrounding water. The water's temperature will decrease as the ice melts. The ice will melt while staying at 0°C. There is a tricky part to solving this problem: When the ice melts it turns into water, and this extra water must be accounted for when calculating the final temperature.

Variables

It is important to distinguish between the two masses of water in the final mixture: the mass that was initially liquid, and the mass that was initially solid ice. We use the subscripts L and S to distinguish these masses of water.

In the section on latent heat, we calculated the heat transferred to the same amount of ice as it melted to be 1.23×10^4 J. Here, we need the heat transferred **from** the water, not **to** the ice. Since the water loses heat, we state this as -1.23×10^4 J, with a negative sign.

mass of liquid water	$m_L = 0.160$ kg
mass of solid ice	$m_S = 0.0370$ kg
heat transferred from water to melt ice	$Q = -1.23 \times 10^4$ J
initial temperature of liquid water	$T_{Li} = 30.0$°C
initial temperature of solid ice	$T_{Si} = 0.00$°C
temperature of liquid water after ice melts	T_{Lf}
temperature of ice-melt	$T_{Sf} = 0.00$°C
final temperature of liquid water plus melted ice	T
heat transferred from liquid water for thermal equilibrium	Q_L
heat transferred to melted ice for thermal equilibrium	Q_S
specific heat of water	$c = 4178$ J/kg·K

What is the strategy?

1. Calculate the temperature of the water after it loses heat to melt the ice, using the specific heat of water.
2. As the water reaches thermal equilibrium, heat transfers **from** the originally liquid water, **to** the melted ice. These are the only heat transfers in the system, so they sum to zero. Use this fact and the specific heat equation to calculate the temperature of the total mass of water at the end.

Physics principles and equations

The specific heat equation

$$Q = cm\Delta T$$

As the two masses of water reach thermal equilibrium, the heat transferred from the originally liquid water plus the heat transferred to the melted ice must sum to zero.

$$Q_L + Q_S = 0$$

Step-by-step solution

First we calculate the temperature of the liquid water after it gives up heat to melt the ice. We use the specific heat of water, 4178 J/kg·K.

Step	Reason
1. $Q = cm\Delta T$	specific heat equation
2. $-1.23\times10^{4}\text{ J} = (4178\text{ J/kg}\cdot\text{K})(0.160\text{ kg})\Delta T$	substitute values
3. $\Delta T = -18.4\text{ K} = -18.4\text{ C°}$	solve
4. $T_{Lf} = T_{Li} + \Delta T$ $T_{Lf} = 30.0°\text{C} + (-18.4°\text{C})$ $T_{Lf} = 21.6°\text{C}$	calculate water temperature

Now we use the fact that the heat transfers sum to zero as the two masses of water reach thermal equilibrium to calculate the final temperature of the total mass of water.

Step	Reason
5. $Q_L + Q_S = 0$	equation above
6. $cm_L(T - T_{Lf}) + cm_S(T - T_{Sf}) = 0$	specific heat equation
7. $T = \dfrac{m_L T_{Lf} + m_S T_{Sf}}{m_L + m_S}$	solve for T
8. $T = \dfrac{(0.160\text{ kg})(21.6°\text{C}) + (0.0370\text{ kg})(0.00°\text{C})}{0.160\text{ kg} + 0.0370\text{ kg}}$	substitute values
9. $T = 17.5°\text{C}$	evaluate

19.21 - Interactive checkpoint: vaporizing mercury

A vial of 0.0500 kg of mercury is at room temperature (20.0°C). What amount of heat must be transferred to the mercury in order to vaporize it? Mercury boils at 357°C, its specific heat is 140 J/kg·K (assume this is constant over the temperature range of interest) and its latent heat of vaporization is 2.95×10^5 J/kg.

Answer:

Q_t = ______ J

19.22 - Conduction

Conduction: The flow of thermal energy directly through a material without motion of the material itself.

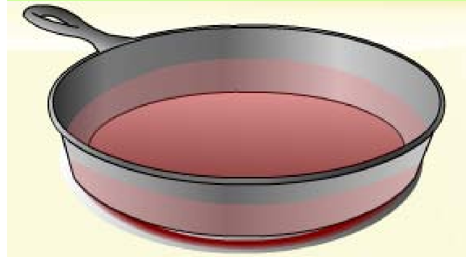
When a frying pan is placed on a burner, heat flows from the burner to the pan. The heat then spreads through the pan, soon reaching the handle even though the handle is not in direct contact with the burner. This process illustrates the flow of thermal energy via conduction.

Conduction is the direct flow of thermal energy without a net motion of the materials involved. The heat flows from the bottom of the pan to its handle. Heat also flows by conduction where the burner and the bottom of the pan are in direct contact.

Conduction results from interactions at the atomic level of particles in an object: molecules, atoms and, in metals, electrons. Particles with greater average energy collide with nearby lower-average-energy particles, increasing their energy. Heat spreads as particles collide with their neighbors, and those neighbors collide with their neighbors, and so on. In this way, heat flows throughout the object. Although motion is involved at this level, the pan itself does not move, and this is part of the definition of conduction: The flow of thermal energy within an object that does not involve motion.

Different materials conduct heat at different rates. This is the topic of the next section.

concept 1



Conduction
Heat flow within object
· Does not involve "bulk" motion

19.23 - Thermal conduction quantified

For appliances that heat or chill (such as ovens or refrigerators), and indeed for an entire house or apartment, thermal conduction is the dominant form of **unintended** heat transfer. Since this transfer of heat is costly and wasteful, a good amount of effort is spent reducing it. In this section, we focus on building materials that are designed to minimize this heat transfer, keeping the interior of a house cool during the summer and warm during the winter.

	Thermal conductivity	Thermal resistance (for 1 inch)	
	k (W/m·K)	*RSI*-value (m²·K/W)	*R*-value (ft²·F°·h/Btu)
Air, sea level (15° C)	0.025	1.00	5.70
Fiberglass (50° C)	0.04	0.64	3.61
Urethane foam	0.06	0.42	2.40
Plywood	0.11	0.23	1.31
Wood (fir)	0.14	0.18	1.03
Water	0.598	0.04	0.24
Concrete (0° C)	0.8	0.03	0.18
Window glass (0° C)	0.95	0.03	0.15
Ice (0° C)	2.14	0.01	0.07
values approximate for building materials			*at 10^5 Pa, 20° C*

Values for *k* and *R*.

The *thermal conductivity* of a material specifies how quickly heat transfers through it. Above, you see a table of thermal conductivity values for materials such as wood, insulating foam, glass, and some metals. The letter k represents the thermal conductivity of a material. Materials with small values for k are used as insulators to reduce the transfer of heat.

The units for thermal conductivity are joules per second-meter-Kelvin. Since watts are joules per second, this is the same as watts per meter-Kelvin, and these are the units used in the table above.

P_c represents the rate of heat transfer through conduction. As shown in Equation 1, P_c is defined as the heat transferred divided by the elapsed time. It is calculated for a slab of material like a glass window pane using the equation shown in Equation 2. The rate of heat transfer increases with the thermal conductivity k of the material that makes up the slab, the area of the slab, and the temperature difference between its two sides. It decreases with the thickness of the material. The equation also can be used to explain why the units for k have meters in the denominator, not "meters squared." Because there is a length term (thickness) in the denominator of the heat-transfer equation, a length term cancels out of the units.

The insulating effectiveness of building materials is usually specified by an R-value. A material with a high R-value is a good insulator. For a slab of material of a given thickness, the *thermal resistance* R is defined, as you see in Equation 3, as the thickness of the slab divided by its thermal conductivity. This is often called the R-value in the building trade. The SI units for R are square meters·kelvin per watt, and the building material ratings are then often called RSI-values. We show metric values with that name in the table above.

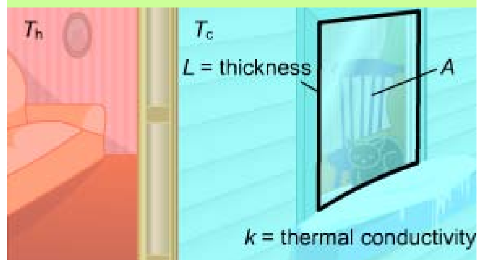
In North America, R-values based on the British measurement system are commonly used for building materials. These R-values have units of square feet·Fahrenheit degrees·hours per British thermal unit. We list them also in the table above. The thickness of the material for both RSI- and R-values is one inch, which is typical for building materials.

The term k in the equation in Equation 2 can be replaced with an expression involving R. This yields another equation, also shown in Equation 3, for the rate of heat transfer.

Considering these equations and studying the table above can help you understand why certain materials are chosen in construction. First, increasing the thickness of a material increases its R-value, so you see thicker insulating material in a colder environment where the temperature difference between inside and outside is greater.

Second, some materials have high k (low R) values, making them unlikely choices for insulators. For instance, building a house out of copper would lead to high heating costs (not to mention building costs). Building materials such as polyurethane foam are

concept 1



Heat conduction
Rate of heat transfer depends on:
· thermal conductivity
· area
· temperature difference
· thickness

equation 1

Rate of heat transfer, definition

$$P_c = Q/t$$

P_c = rate of heat transfer (J/s)
Q = heat transferred
t = time

equation 2

Rate of heat transfer, calculated

effective insulators and can be combined with other reasonably good insulators such as wood for even greater energy efficiency.

Third, materials can be combined. Double-paned windows trap a quantity of an inert gas like argon between two layers of glass. Argon has a high R value and considerably reduces the rate of heat transfer through the window.

$$P_c = \frac{kA\Delta T}{L}$$

k = thermal conductivity
A = area
ΔT = temperature difference
L = thickness
k units: J/s·m·K = W/m·K

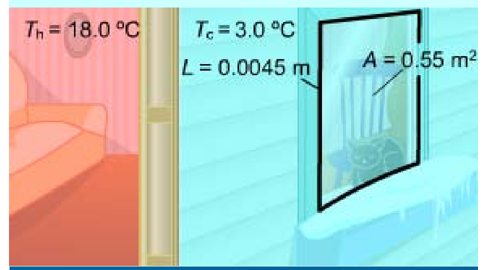
equation 3

Thermal resistance

$$R = L/k$$
$$P_c = \frac{A\Delta T}{R}$$

R = thermal resistance
R units (SI): m^2·K/W
R units (British): ft^2·F°·h/Btu

example 1



Heat transfers through the window at a rate of 1700 J/s. What is its thermal conductivity constant?

$$P_c = \frac{kA\Delta T}{L}$$

$$k = \frac{P_c L}{A\Delta T}$$

$$k = \frac{(1700\text{ J/s})(0.0045\text{ m})}{(0.55\text{ m}^2)(18.0°\text{C} - 3.0°\text{C})}$$

k = 0.93 W/m·K

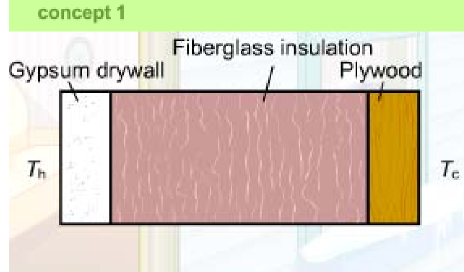
19.24 - Conduction through composite objects

Real-world objects such as the walls of a house are often a composite of different materials. For example, a house wall may consist of gypsum drywall, fiberglass insulation and plywood. At the right, you see a schematic of a wall made of materials of varying thicknesses.

To calculate the rate of heat flow through this composite object, the overall thermal resistance is calculated by summing the resistance of each object. This value can be used as the R-value of a single object in other equations. You see this in Equation 1 on the right.

When designing buildings, the rate at which heat will flow through the walls is an important consideration. Example 1 shows a calculation of the rate of heat flow using R-values for three common building materials.

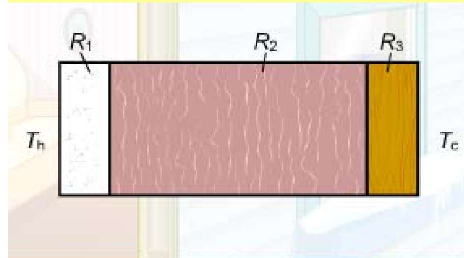
concept 1



To calculate rate of heat transfer:

Add thermal resistance of each layer

equation 1

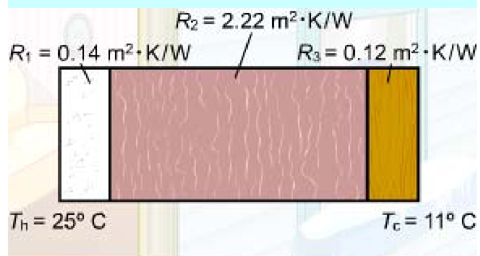


To calculate rate of heat transfer:

$$P_c = \frac{A\Delta T}{R_{comp}}$$

P_c = overall heat transfer rate
A = area
ΔT = temperature difference
$R_{comp} = R_1 + R_2 + \ldots + R_n$
R_n = thermal resistance of layer n

example 1



A wall of a house is constructed as shown. What is the rate of heat transfer though a 12.0 m^2 area of the wall?

$R_{comp} = R_1 + R_2 + R_3$
$R_{comp} = (0.14 + 2.22 + 0.12)\ \mathrm{m^2{\cdot}K/W}$
$R_{comp} = 2.48\ \mathrm{m^2{\cdot}K/W}$

$$P_c = \frac{A\Delta T}{R_{comp}}$$

$$P_c = \frac{(12.0\ \mathrm{m^2})(25\ \mathrm{°C} - 11\ \mathrm{°C})}{2.48\ \mathrm{m^2{\cdot}K/W}}$$

$P_c = 68\ \mathrm{J/s}$

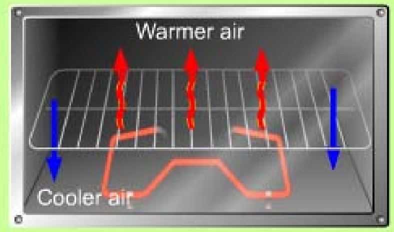
19.25 - Convection

Convection: Heat transfer through a gas or liquid caused by movement of the fluid.

Gases and liquids usually decrease in density when they are heated (liquid water near $0°C$ is a notable exception). When part of a body of liquid or gas is heated, the warmed component rises because of its decreased density, while the cooler part sinks. This occurs in homes, where heat sources near the floor heat the nearby air, which rises and moves throughout the room. The warmer air displaces cooler air near the ceiling, causing it to move near the heat source, where it is heated in turn. This transfer of heat by the movement of a gas or liquid is called *convection*.

All kitchen ovens, like the one shown in Concept 1, rely largely on convection for baking. The heating element at the bottom of the oven warms the air next to it, causing it to rise. The heated air then reaches the food in the oven to warm it, while the cooler air sinks to the bottom of the oven. So-called "convection ovens" speed this process with fans that cause the air to circulate more quickly.

concept 1



Convection
Heat transfer due to movement in gases and liquids

Convection occurs in liquids as well as in gases. If you stir spaghetti sauce as it heats, you are accelerating the process of convection. Again, your goal is to uniformly distribute the thermal energy.

If you see a hawk soaring upward without flapping its wings, it may be riding what is called a "thermal." As the Sun warms the ground, the nearby air also becomes warmer. In the process, it becomes less dense, and is forced upward by air that is cooler and denser. A bird can ride this upward draft.

19.26 - Radiation

Radiation: Heat transfer by electromagnetic waves.

If you place your hand near a red-hot heating element and feel your hand warm up, you are experiencing thermal radiation: the transfer of energy by electromagnetic waves. You correctly think of objects like the heating element as radiating heat; in fact, every object with a temperature above absolute zero radiates energy.

Radiation consists of electromagnetic waves, which are made up of electric and magnetic fields. Radiation needs no medium in which to travel; it can move through a vacuum. The wavelength of radiation varies. For instance, red light has a wavelength of about 700 nm, and blue light a wavelength of about 500 nm. Infrared and ultraviolet radiation are two forms of radiation whose wavelengths are, respectively, longer and shorter than those of visible light.

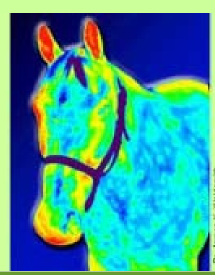
concept 1

Radiation
Heat transfer by electromagnetic radiation

All objects radiate electromagnetic radiation of different wavelengths. For instance, you see the red-hot stove coil because it emits some visible light. The coil also emits infrared radiation that you cannot see but do feel as heat flowing to your hand, and it emits a minimal amount of ultraviolet radiation too.

Although any particular object radiates a range of wavelengths, there is a peak in that range, a wavelength at which the power output is the greatest. This peak moves to shorter wavelengths as the temperature of the object increases. Understanding the exact form of the spectrum of thermal radiation wavelengths requires concepts from quantum physics, and its derivation was one of the early triumphs and verifications of quantum theory.

concept 2



Radiation
All objects emit radiant energy

Bodies with temperatures near the temperature of the surface of the Earth emit mostly infrared radiation. In the photograph in Concept 2, called an *infrared thermograph*, you see the radiation emitted by a horse. Since areas of inflammation in the body are unusually warm, and emit extra thermal radiation, veterinarians can use photographs like this to diagnose an animal's ailments. They are created by a digital or film camera that assigns different (visible) colors to different intensities of (invisible) infrared radiation in a process called false color reproduction.

Sunlight is a form of radiation and is crucial to life on Earth. The Sun emits massive amounts of energy in the form of radiation: 3.9×10^{26} joules every second. Some of that strikes the Earth, where it warms the planet and supplies the energy that plants use in photosynthesis.

The amount of power radiated by a body is proportional to the fourth power of its absolute temperature, its surface area, and a factor called its emissivity. The Sun emits tremendous amounts of radiation energy because it is quite hot (about 6000 K) and vast (with a surface area of about 6×10^{18} m^2).Only a small portion of the total power emitted by the Sun reaches the Earth. Even this fraction is an enormous amount: 1.8×10^{17} watts, about 100 times what human civilization consumes. The average solar power striking the Earth's atmosphere in regions directly facing the Sun is about 1370 watts per square meter. This value is called the *solar constant*.

Not all parts of the Earth directly face the Sun, and some radiation is reflected or scattered by the Earth's atmosphere before it reaches the surface. Different regions on Earth receive different average amounts of power per square meter: Measurements show that on average about 240 W/m^2 (watts per square meter) reaches the Earth's surface. North America is estimated to receive radiation of 150 W/m^2, on the average.

Given the dimensions of an average American house, this amount of radiation supplies about four times as much power as the household consumes. Using this clean "renewable resource" constructively for human purposes challenges both engineers and physicists. It also provides a significant opportunity to conserve non-renewable energy resources that would otherwise be used to heat, cool, and light our houses.

Efforts have been underway to take advantage of this form of energy for years. *Photovoltaic cells* convert solar energy directly into electricity.

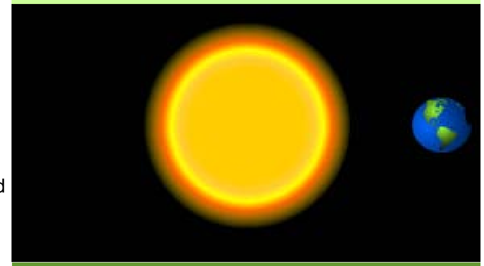
The energy in sunlight can also be used without transforming it into electricity. The heat of sunlight can be used to heat both water and the interior of a house. Devices called *solar collectors* are becoming increasingly common. They are typically black; you see one in Concept 4. Solar collectors heat water for uses ranging from hot showers to warm swimming pools. Typically, the collectors contain tubes or rods that hold water, and after it is heated the water is circulated to the desired locations in the building.

Current solar collectors come in a variety of shapes and sizes. The engineers designing them have two main challenges. One is to maximize the amount of solar radiation that strikes the collector. Some collectors, particularly those designed for industrial or research purposes, use mirrors to focus sunlight onto the water tubes, increasing the amount of solar radiation that reaches the tubes.

The second engineering challenge is to efficiently transfer the heat energy of sunlight to the water inside the collector tubes and then circulate the heat out of the collector. There are many designs intended to maximize the efficiency of this process.

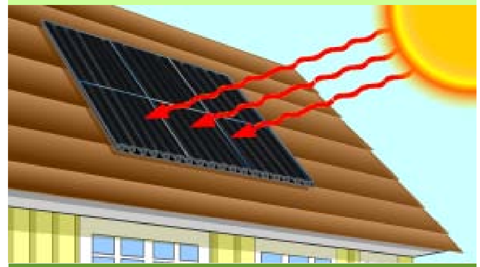
Residential systems can capture usable amounts of energy. One current system is capable of capturing 1000 Btu (British thermal units) per square foot of collector per day. A typical residential gas or oil furnace can supply on the order of 1.2 million Btu a day, which means a 10-by-20 foot solar panel supplies about one-sixth of the energy of a typical home furnace.

concept 3



Sun and Earth
Sun emits enormous amounts of radiant energy
Some strikes the Earth's atmosphere

concept 4



Solar collector
Uses solar energy that reaches the Earth

If you are interested in solar collectors, you can view a web site with information on them. You will study the physics of electromagnetic radiation in detail in a later chapter.

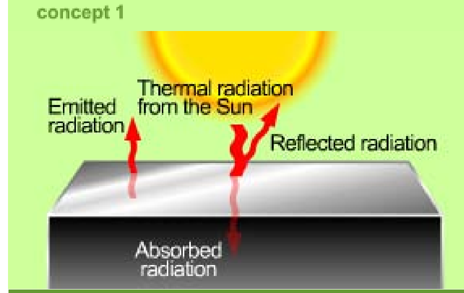
19.27 - Radiation quantified

When radiation reaches the surface of an object, it is either absorbed or reflected. Any ordinary object absorbs some of the incident radiation, and reflects some back. A material's *emissivity* is the ratio of absorbed radiation to incident radiation.

An object called a *blackbody* will absorb all the radiation that strikes it. No real object is a perfect blackbody, but the ideal object is the theoretical limit of bodies that absorb almost all the radiation that hits them. A blackbody has an emissivity of one because it absorbs all radiation. A material that reflected all radiation would have an emissivity of zero. (Emissivity varies by the wavelength of the electromagnetic radiation striking a body, so the discussion in this section applies to the **average** emissivity across all wavelengths.) At the right, you see a table of emissivity values for various materials.

All objects above absolute zero in temperature emit radiation. It can be shown that the amount of radiation absorbed by a body in thermal equilibrium with its surroundings equals the radiation it emits (which is why the proportion of radiation absorbed can reasonably be called "emissivity"). This equality means that not only does a blackbody represent the theoretical limit for the maximum amount of energy absorbed, but also the theoretical limit of the radiation emission.

concept 1



Radiation
Objects absorb, reflect electromagnetic radiation
Objects also emit radiation

At first empirically, and later through theory, it has been demonstrated that the rate at which a body emits radiation is a product of four factors: the fourth power of its Kelvin temperature, its surface area, the *Stefan-Boltzmann* constant σ, and the emissivity constant for the material. This is the first equation shown in Equation 1. The rate at which a body absorbs radiation is calculated similarly, but uses the temperature of the surrounding environment. You see this as the second equation.

About 30% of the sunlight hitting the Earth is reflected back into space (mostly by clouds). In Example 1 we ignore the reflection of sunlight and assume that the Earth behaves like a blackbody. We use the solar constant and Equation 1 to provide an approximation of what the temperature near Earth's surface should be. The result is 279 K $(6°\text{C})$, slightly lower than the actual average temperature of the atmosphere close to the Earth's surface, which is about 288 K $(15°\text{C})$. If we took into account the energy reflected by the Earth into space, and the differing emissivities of the Earth for the reflected and emitted wavelengths, the calculated temperature would be a rather chilly 255 K $(-18°\text{C})$.

The Earth is not a perfect blackbody radiator, of course, but its surface temperature is still greater than the radiation equation would predict.

This example serves to illustrate the role of the greenhouse effect. The Earth's atmosphere "traps" a substantial amount of the radiation emitted by the Earth's surface. Without this effect, the temperature at the surface of the Earth would be cooler.

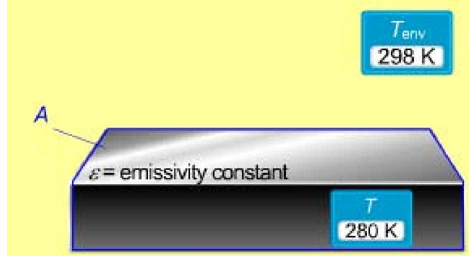
concept 2

	Emissivity
Polished silver, gold, aluminum	0.02-0.04
Mercury (the element)	0.09-0.12
Venus	0.24
Earth average	0.67
White enamel paint	0.87-0.91
Mercury (the planet)	0.9
Flat black lacquer	0.92-0.96
Candle soot	0.95

Emissivity

ε = ratio absorbed/incident radiation

equation 1

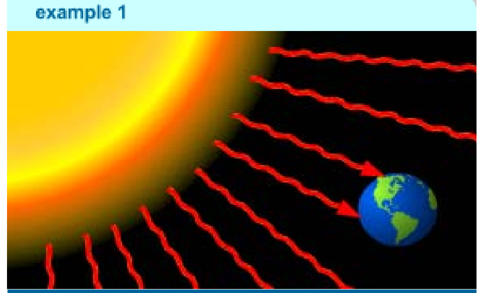


Radiated and absorbed power

$$P_{rad} = \sigma \varepsilon A T^4$$
$$P_{abs} = \sigma \varepsilon A T_{env}^{\ 4}$$

P_{rad} = power radiated
P_{abs} = power absorbed
ε emissivity, A = surface area
T = temperature of object (K)
T_{env} = temperature of environment (K)
$\sigma = 5.67\times10^{-8}$ W/m^2·K^4
(Stefan-Boltzmann constant)

example 1



Estimate the Earth’s equilibrium surface temperature by modeling the Earth as a blackbody. In this idealization the Earth does not reflect any sunlight, so the intensity of the incident sunlight equals the solar constant, 1370 W/m^2.

$$P_{in} = P_{out}$$
$$I\left(\pi R_e^2\right) = \sigma \varepsilon T^4 \left(4\pi R_e^2\right)$$
$$T = \sqrt[4]{\frac{I}{4\sigma\varepsilon}}$$
$$T = \sqrt[4]{\frac{1370\ \text{W/m}^2}{4\left(5.67\times10^{-8}\ \text{W/m}^2\cdot\text{K}^2\right)(1)}}$$
$$T = 279\ \text{K}$$

19.28 - Global warming and the greenhouse effect

The topic of global warming is frequently in the news and inspires much discussion. The debate can be extraordinarily passionate at times, because if humans are truly accelerating the increase in the average temperature of the planet, the consequences may prove severe.

A mechanism called the *greenhouse effect* plays a significant role in determining the temperature of the Earth's atmosphere, as it does for other planets with atmospheres. This effect was discovered in the 19th century and is well understood.

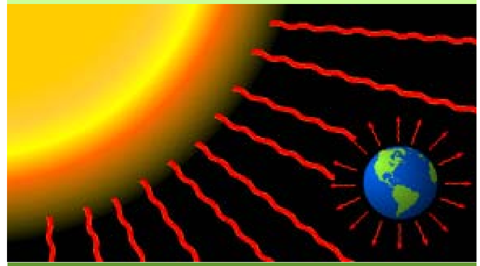
The Earth absorbs energy from the Sun. In turn, the Earth radiates energy, primarily in the infrared spectrum. We measure both in terms of power: the rate at which energy is absorbed or emitted. The rate of radiation emission is proportional to the Earth's temperature to the fourth power and to the Earth's surface area. Using the Earth's surface area and its average temperature (about 288 K), the power radiated by the Earth can be determined.

The power of all the Sun's radiation that reaches the Earth must equal the power of the radiation emitted by the Earth. If the power of solar radiation were to decrease, then the Earth would absorb less energy over time, it would cool and its rate of emission would decrease. Equilibrium of power absorbed and radiated would be reached at a lower temperature. The converse is also true: If the power of the Earth's emission falls below the power of the radiation it absorbs, the planet warms up, the power of its emission increases, and a new equilibrium is reached.

The Earth receives about 240 watts per square meter of thermal radiation from the Sun. You could use the radiation equation to calculate the power of its emitted radiation due to its observed average temperature of 288 kelvins, its emissivity and its surface area. The result would be greater than the power of the radiation it absorbs, which means the Earth should be cooling. The equilibrium temperature of the Earth should be a rather chilly 255 K.

The greenhouse effect reconciles this imbalance. The effect was noted in (surprise) a greenhouse by French scientist Jean-Baptiste Fourier in 1827. Visible light enters the greenhouse and it warms objects inside the greenhouse, which then re-emit energy as infrared radiation. Fourier noted that greenhouses are partly kept warm because the glass absorbs most of the infrared radiation emitted by the contents of the greenhouse as they are warmed by the heat from the Sun. The glass then re-emits that radiation, and since it does so in all directions, some of the radiation is directed back inside the greenhouse.

concept 1



Solar, terrestrial radiation
Sun's radiation warms Earth
Earth emits radiation

concept 2

Greenhouse effect
Atmosphere absorbs some radiation from Earth
· Re-emits radiation in all directions
· Radiation toward Earth warms planet
Mix of gases alters temperature of atmosphere

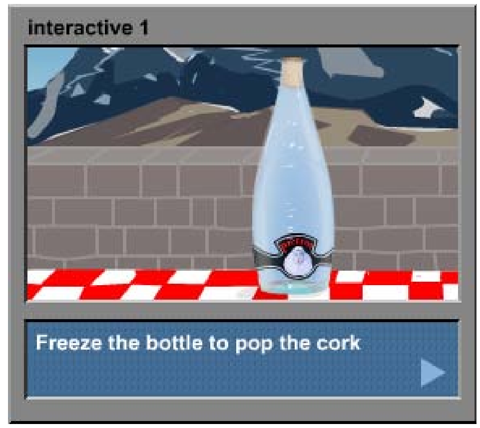
The Earth's atmosphere has the same effect: It absorbs infrared radiation emitted by the Earth and then re-emits some of it back toward the Earth as illustrated in Concept 2. This accounts for the missing 33 kelvins (288 K − 255 K) mentioned above. The Earth is kept warmer by its atmosphere. This effect is even more pronounced on some other planets, among them Venus, which have an atmospheric composition richer in the so-called "greenhouse gases."

The greenhouse effect has been operating on the Earth for billions of years. It is not caused by humans. As the composition of the Earth's atmosphere has changed over time, so has the magnitude of this natural greenhouse effect.

In the past 50 years there has been increasing concern about the degree to which human activities have altered the composition of the atmosphere. Since 1957, there have been a series of studies showing that the amount of carbon dioxide in the atmosphere has been steadily and rapidly increasing. In addition to natural processes (such as animal respiration), burning coal and oil produces vast amounts of carbon dioxide. Carbon dioxide is a *greenhouse gas*, especially efficient at absorbing and re-emitting infrared radiation.

The Earth absorbs some of this gas (for instance, the process of photosynthesis in forests consumes some fossil fuel carbon dioxide, as does the global ocean) but the current level of production far exceeds the Earth's ability to absorb it. Scientists estimate that at current rates the amount of carbon dioxide in the Earth's atmosphere will double in less than a century, perhaps by 2050. This leads to what is called the *enhanced greenhouse effect*. In this case, enhanced does not mean "improved."

19.29 - Interactive summary problem: pop the cork



You just bought a bottle of Pierrot, the water from ancient limestone caves deep in the French Alps, filtered by pure quartz crystals. But you did not realize the bottle came with a cork, and you have no corkscrew. Fortunately, your knowledge of thermal physics comes in handy.

You remember that the density of ice is 9% less than the density of water. This means that water expands quite a bit when it freezes into ice. If you let the water in the bottle freeze, the expansion of the ice will push the cork out.

If 89.0% of the water freezes, the expanding ice will just push the cork out. But if more than 89.0% of the water freezes, the ice will expand too much and the bottle will break. You want to remove just enough heat from the water so that exactly 89.0% of it turns to ice.

In the interactive simulation on the right, you control the amount of heat removed from the water. The bottle contains 0.750 kg of water and its temperature is now $15.0°C$. You need to remove enough heat to reduce the temperature of all the water to $0°C$, and then remove enough additional heat to freeze 89.0% of it. To do these calculations, you will need to use the specific heat of water, 4178 J/kg·K, and the latent heat of fusion of water, $3.34×10^5$ J/kg. We ignore the glass bottle itself in these calculations. Heat is removed from the bottle, but much more heat (about 50 times more) is removed from the water. Also, while the volume of the glass bottle decreases slightly as it cools, the expansion of the ice is much greater. Similarly, the small air space at the top of the bottle has little effect.

Set the amount of heat to be removed from the water, then press GO. If you are right, the ice will push the cork out. Press RESET to try again.

If you have trouble calculating the correct amount of heat transfer, review the sections in this chapter on specific heat and latent heat, and the sample problem that combines the two concepts.

19.30 - Gotchas

Heat is the same as temperature. No, heat is a flow of energy. Temperature is a property of an object. The flow of heat will change the temperature of an object, and a thermometer measures the object's temperature.

The Fahrenheit temperature system is the wave of the future. If you think so, can I interest you in buying a record player?

Two rods of the same material experience the same increase in temperature, which means they must have expanded by the same amount of length. Only if they were the same initial length. Their percentage increase would be the same in any case.

You throw a football upward. You have not increased the internal energy of the air within the football. Correct: You have not increased the internal energy of the molecules inside the football. (You have increased its translational kinetic energy and its rotational kinetic energy and its gravitational PE by throwing it upward, but not its internal energy.)

19.31 - Summary

Temperature can be thought of simply as what thermometers measure. Thermometers rely on physical properties that vary with temperature in a reliable, reproducible way.

Different scales are used to measure temperature. The Fahrenheit scale is commonly used to measure temperature in the United States, but the Celsius scale is used in most of the rest of the world, and the Kelvin scale is used in science. Reference points such as the freezing and boiling points of water define a temperature scale and how to convert between scales. Another important (though ultimately unreachable) temperature point is absolute zero, where molecules are at their minimum energy state. This is 0 K on the Kelvin scale, or $-273.15°C$.

A constant-volume gas thermometer relies on the fact that the pressure of a gas varies linearly with its temperature. Such a thermometer can measure temperatures very precisely.

Heat is the flow of energy between objects resulting from a difference in temperature. Like work, heat changes the energy of an object, but we cannot say that an object *has* a certain amount of heat, any more than we could say it contains a certain amount of work. Heat can change the internal energy of a system. Heat is commonly measured in joules, or sometimes in calories or British Thermal Units. A food Calorie is actually a kilocalorie, or 1000 calories.

The zeroth law of thermodynamics states that two objects, each in thermal equilibrium with a third object, would also be in thermal equilibrium with each other if they were placed in direct contact.

Objects change in length and volume when their temperature changes, in a process called thermal expansion. The change in length of an object depends on the coefficient of linear expansion of the material from which it is made, which is represented by the symbol α. Its change in volume depends on the coefficient of volume expansion, represented by β. Water has the unusual property that it expands with decreasing temperature from $4°C$ to $0°C$; this is why ice floats.

The heat capacity of an object determines the amount of heat required to produce a given change in its temperature.

Specific heat is a property of a material, not an object, that relates the amount of heat to the change in temperature for a unit of mass of the material.

Molar specific heat is similar, but relates heat to temperature change per mole of a material. A mole is 6.022×10^{23} particles (typically molecules) of a material.

The transformation of an object from one state (solid, liquid, or gas) to another is a phase change. Phase changes require heat to be either added or taken away. The latent heat of a material is the energy consumed or released per kilogram during a particular phase change. The latent heat of fusion is the heat flow during a change between solid and liquid; the latent heat of vaporization is the heat flow during a change between liquid and gas.

Ways in which heat is transferred include conduction, convection, and radiation. Conduction is the flow of thermal energy directly through a material without motion of the material itself. Thermal resistance, sometimes called the R-value, determines the rate of heat transfer by conduction. Convection transfers heat by the bulk movement of molecules in a gas or liquid. Radiation transfers energy by means of electromagnetic waves. Every object emits electromagnetic radiation.

Emissivity is the ratio of the absorbed to incident radiation for an object. An object with an emissivity of one is called a blackbody; it absorbs all the radiation that strikes it, and is in turn a perfect radiator when it is in thermal equilibrium with its surroundings.

Equations

$$T_K = T_C + 273.15$$

$$T_C = (5/9)(T_F - 32)$$

Thermal expansion

$$\Delta L = L_i \alpha \Delta T$$

$$\Delta V = V_i \beta \Delta T$$

Specific heat

$$Q = cm\Delta T$$

Molar specific heat

$$Q = kn\Delta T$$

Latent heat

$$Q = L_f m$$

$$Q = L_v m$$

Thermal conduction

$$P_c = Q/t$$

$$R = L/k$$

$$P_c = \frac{kA\Delta T}{L} = \frac{A\Delta T}{R}$$

$$P_c = \frac{A\Delta T}{R_{comp}}$$

Radiation

$$P_{rad} = \sigma\varepsilon A T^4$$

$$P_{abs} = \sigma\varepsilon A T_{env}^4$$

Chapter 19 Problems

Conceptual Problems

C.1 The standard value for a healthy human's body temperature is $37°C$, to the nearest Celsius degree, although this value is only an approximation. In fact, the temperature of a healthy person varies by as much as $1.5°C$ each day, being lowest in the morning and peaking in the evening. On North American fever thermometers, a "normal" body temperature is usually indicated as $98.6°F$. (a) Explain how this "normal" value in the Fahrenheit system is obtained from the "normal" value in the Celsius system. (b) Explain why stating the value for a "normal" human temperature with three significant digits, as $98.6°F$, is misleading. (c) What would be a better "normal" value in the Fahrenheit system?

(a)
(b)
(c) __________ ° F

C.2 A new website comes out promoting the Cold Water Diet. It proclaims: "On the CWD, you eat all the carbs and fat you want, and just drink cold water. One liter of cold water has a mass of 1000 g, and it takes 1.0 calorie to raise the temperature of 1.0 g of water by 1 Celsius degree. If you ingest a liter of water at 0°C, it will take 37,000 calories to raise the temperature up to 37°C, which is body temperature. That energy has to come from your body, so you can watch those pounds melt away!" What is wrong with the website's claim?

C.3 A steel block is heated so that the length of each side increases 1%. What happens to its mass?

i. It increases 1%
ii. It increases 3%
iii. It increases 3.0301%
iv. It does not change

C.4 You want to use a calorimeter to measure the specific heat of a very small object that you suspect has a low specific heat value. Assume that you can use any of three liquids to fill the calorimeter: water, ethyl alcohol (which has a specific heat of 2549 J/kg·K), or mercury (which has a specific heat of 139 J/kg·K). Which liquid would you choose, and why?

○ Water ○ Ethyl alcohol ○ Mercury

C.5 Why does a tiled floor feel colder to your bare feet than a carpeted floor, even though they are both at room temperature?

C.6 In coastal towns, breezes tend to blow offshore (from the land to the ocean) during the night and early morning, and onshore (from the ocean to the land) during the afternoon and evening. What factors primarily contribute to this effect?

☐ Heat conduction
☐ Heat convection
☐ Specific heat of land vs water
☐ Salinity of the water
☐ Thermal expansion

C.7 In outer space environments, such as on the International Space Station, why must every piece of machinery or electronic equipment have its own cooling fan? This makes for a noisy environment. Hint: What is different compared to an Earth-based laboratory, and why should this matter?

C.8 When people are extremely cold and are not prepared with warm clothing, they will sometimes crouch and curl up into a tight roll, hugging their knees to their chests. How does such a position conserve body heat?

Section Problems

Section 2 - Temperature scales

2.1 Daniel Fahrenheit first proposed the temperature scale that bears his name in 1724. Originally, his reference points were a well-mixed slurry melted from equal weights of ice and salt (similar to what you use in a hand-cranked ice cream freezer) for 0 degrees, and the temperature of the healthy human body for 12 degrees. Later he subdivided each one of his original "degrees" into eight equal parts to define degrees Fahrenheit. In modern degrees Fahrenheit, what is the temperature of (a) the ice-salt mixture? (b) Fahrenheit's "healthy human"?

(a) __________ ° F
(b) __________ ° F

Section 3 - Temperature scale conversions

3.1 Convert each of the following temperatures to the indicated scale: (a) 100.00 K to the Celsius scale; (b) 100.00 K to the Fahrenheit scale; (c) 72°F to the Celsius scale; and (d) −273.15°C to the Fahrenheit scale.

(a) __________ ° C
(b) __________ ° F
(c) __________ ° C
(d) __________ ° F

3.2 What Celsius temperature is the same when converted to Fahrenheit? (There is only one such temperature.)

__________ °

3.3 An iron rod is heated from 66.0°F to 280°F. Express the increase in temperature in Celsius.

__________ ° C

3.4 In 1731 the French scientist Rene Reaumur proposed the Reaumur temperature scale. The reference points for his scale were the freezing and boiling points of water (0°R and 80.0°R respectively). Like the Fahrenheit scale in northern Europe, his scale was widely used in France until the 1960's, when the Celsius scale was officially adopted. (a) Convert 37.0°C to Reaumur. (b) Convert 56.5°F to Reaumur. (c) Convert 256 K to Reaumur.

(a) __________ ° R
(b) __________ ° R
(c) __________ ° R

Section 9 - Thermal expansion: linear

9.1 The plumbing in an old house uses lead pipes (which are now considered hazardous). Hot water is run through a section of pipe, increasing its temperature from 19.4°C to 36.8°C. If the pipe is initially 3.90 m in length, what is the change in the length of the pipe?

__________ m

9.2 A square copper plate has a circular hole in its center. Each side of the plate is 23.52 cm, and the radius of the circular hole is 7.827 cm. If the plate's temperature increases 174.3 K, what are the new lengths of (a) the side of the plate and (b) the radius of the hole? Express your answers to four significant digits.

(a) __________ cm
(b) __________ cm

9.3 A metal rod is 2.673 m long at 23.25°C. When its temperature is increased to 168.4°C, the length of the rod is 2.681 m. What is the metal's coefficient of linear expansion? Express your answer to four significant digits.

__________ 1/C °

9.4 The aptly-named Steel Bridge over the Willamette River in Portland, Oregon, has a central span steel truss that is 64.1 m long. What is its change in length over the year from the average minimum annual temperature of 0.90°C to the average maximum of 26.3°C? Use a coefficient of linear expansion of 1.17×10^{-5} (1/C°).

__________ m

9.5 A ring of silver has an inside diameter of 2.355 cm at a temperature of 23.48°C. A steel rod has a diameter of 2.361 cm. To what temperature should you raise the ring so it will just slip onto the rod? Assume the rod does not change temperature. Express your answer to four significant digits.

__________ ° C

9.6 An aluminum ring has an inside diameter of 1.793 cm, and an iron rod has a diameter of 1.798 cm, both at 18.73°C. If you increase the temperature of both the ring and the rod, at what common temperature will the ring just slip onto the rod?

__________ ° C

9.7 An aluminum beam shows a length of L meters as measured with a steel measuring tape, with both the beam and the tape at a temperature T_i. The temperature (of both the beam and the measuring tape) then increases by ΔT. Develop an expression for the value the tape now shows for the length of the beam. Hint: Find out how long a meter becomes on the steel measuring tape, find the new length of the aluminum beam, then think about how to combine the two results.

9.8 A clock is designed with a simple pendulum, consisting of a mass at the end of an aluminum rod, whose period is supposed to be equal to one second. The clock is designed to keep accurate time at 20.0°C, but is operated at a constant temperature of −80.0°C. (a) Will the clock run fast or slow? Explain. (b) How many periods will the pendulum go through before the clock is off by 1.00 s? (The coefficient of linear expansion of aluminum is 2.31×10^{-5} 1/C°.)

(a) ◯ Fast ◯ Slow
(b) __________ periods

Section 10 - Sample problem: thermal expansion and stress

10.1 A copper rod at a temperature of 22.82°C is bolted in place between two rigid walls. What stress does the rod exert (the force on each of the perfectly rigid walls, divided by the area of contact) when it is heated to a temperature of 274.2°C? (The coefficient of linear expansion for copper is 1.65×10^{-5} 1/C°, and Young's modulus for the material is 1.08×10^{11} N/m^2.)

__________ N/m^2

10.2 A rod has a cross-sectional area of 0.001615 m^2. When the rod's temperature is 12.23°C, it is placed between two rigid walls that are exactly as far apart as the length of the rod. The rod's temperature then increases to 27.23°C, and the rod exerts a force of 5.334×10^4 N on each wall. If Young's modulus for the material of which the rod is made is 1.05×10^{11} N/m^2, what is the coefficient of linear expansion of the rod?

__________ 1/C °

Section 12 - Thermal expansion: volume

12.1 A lead ball has a volume of 94.3 cm^3 at 19.3°C. What is the change in volume when its temperature changes to 32.3°C?

__________ cm^3

12.2 A solid copper ball with radius 1.35 cm increases in temperature from 15°C to 86°C. What is the change in its volume?

__________ cm^3

12.3 A hot water tank holds exactly 80.0 gallons of water at 115°F. What is the volume of the water at 40.0°F?

__________ gallons

12.4 The density of mercury at 0°C is 13,595 kg/m^3. What is its density at 100.0°C?

__________ kg/m^3

12.5 Fever thermometers are often made of a glass tube filled with a liquid. Older thermometers were often filled with mercury. Newer ones are filled with red-tinted ethyl alcohol. You have two identical glass tube thermometers, one filled with mercury and the other filled with ethyl alcohol. The degree marks on the mercury thermometer are 1.32×10^{-3} m apart. How far apart are they on the alcohol thermometer? (Ignore the expansion of the glass in your calculations. The coefficient of volume expansion is 1.12×10^{-3} 1/C° for ethyl alcohol.)

__________ m

12.6 A cup is made from a thin sheet of aluminum, shaped into a cylinder open at one end. At 13.4°C, the radius of the cylinder is 3.24 cm and its height is 7.56 cm. (a) What is the change in volume of the cup when it is raised to a temperature of 78.2°C? (The coefficient of linear expansion for aluminum is 2.31×10^{-5} 1/C°.) (b) If the cup is initially exactly filled to the brim with water, and the water is raised to the same temperature as the cup, what volume of water bulges out over the top of the cup? (The coefficient of volume expansion for water is 20.7×10^{-5} 1/C°.)

(a) __________ cm^3
(b) __________ cm^3

12.7 The temperature of a block of lead is raised from $0°\text{C}$ to $100°\text{C}$. What is the percentage change in its density? (The density of lead at $0°\text{C}$ is $11{,}300\ \text{kg/m}^3$.)

__________ %

12.8 Prove that for a solid cube of any material, the coefficient of volume expansion is approximately three times the coefficient of linear expansion.

Section 14 - Specific heat

14.1 A block of iron at 35.0°C has mass 2.30 kg. If 3.50×10^4 J of heat are transferred to the block, what is its resulting temperature?

__________ ° C

14.2 You want to take a bath with the water temperature at 35.0°C. The water temperature is 38.0°C from the hot water tap and 11.0°C from the cold water tap. You fill the tub with a total of 191 kg of water. How many kilograms of water from the hot water tap do you use?

__________ kg

14.3 An orchid grower hears that the temperature is expected to fall below freezing overnight, so she puts a large barrel in her orchid greenhouse and fills it with 315 kg of water at 21.0°C. In the morning, the water is at 1.5°C. How much heat left the water during the night? (Express the answer as a positive amount.)

__________ J

14.4 A 1.32 kg sample of some material increases in temperature from 13.5°C to 39.2°C when 1.98×10^4 J of heat energy is transferred to it. What is the specific heat of the material?

__________ J/kg · K

14.5 A 2.4 kg iron ball is dropped from a height of 14 m onto a concrete roadway, and 2.5% of its kinetic energy at the time it reaches the ground is transformed into internal energy in the ball itself. (The rest of the energy is transmitted to the ground, converted into sound energy, and so on.) What is the ball's increase in temperature?

__________ K

14.6 A 0.00340 kg lead bullet is traveling at 473 m/s when it hits the bull's-eye of a practice target. Half of the bullet's kinetic energy is transferred to the bullet in the form of heat. What is the resulting temperature increase in the bullet? Assume that the specific heat of the lead does not vary significantly over this temperature.

__________ K

14.7 A lead block and a silver block are both at the temperature of 41°C. They then each absorb the same amount of heat, and, without any exchange of heat between them or with the environment, their new final temperatures are again the same. (a) If the mass of the lead block is 5.0 kg, what is the mass of the silver block? (b) If they each absorbed 6.9×10^3 J, what is their final temperature?

(a) __________ kg
(b) __________ ° C

14.8 As a body of water gives up heat energy to the air above it, the air temperature increases. Assume 1.0 kg of water (about one liter) decreases 1.0 K in temperature, giving up its energy to a nearby body of air. (a) What mass of air could the energy from the water increase by 1.0 K in temperature? (b) Assume the density of the air is 1.2 kg/m^3. What is the volume of this mass of air?

(a) __________ kg
(b) __________ m^3

14.9 A samovar is a Russian urn, often made of silver, for serving tea. A large samovar made from 4.62 kg of silver is filled with 11.3 kg of tea, at the same temperature. How much heat is required to raise the temperature of both the samovar and tea by 34.7 C°? (Assume the specific heat of tea is the same as that of water, 4178 J/kg·K.)

__________ J

Section 15 - Sample problem: a calorimeter

15.1 A large handful of identical gemstones, totaling 195 carats (one carat has a mass of 200 mg) at temperature 73.2°C, is placed in a calorimeter filled with 0.594 kg of water at 13.6°C. After thermal equilibrium is established, the water temperature is 14.1°C. What is the specific heat of the gemstone material?

__________ J/kg · K

15.2 An Italian chef, who is very scientific in her approach to cooking, measures the specific heat of a porcini mushroom by placing it in a calorimeter filled with 0.678 kg of olive oil at 13.2°C. The mushroom has mass 0.0802 kg and temperature 24.3°C before being placed into the calorimeter. The olive oil's temperature is 14.7°C after thermal equilibrium is reached. The specific heat of olive oil is 1970 J/kg·K. What is the specific heat of the mushroom?

__________ J/kg · K

Section 17 - Molar specific heat

17.1 A lump of copper increases in temperature 4.5 K when 2.4×10^2 J of heat energy is transferred to it. How many moles of copper are in the lump?

__________ mol

Section 19 - Latent heat

19.1 An ice cube with a mass of 0.0410 kg at 0°C melts to water with no change in temperature. How much heat does the ice absorb while melting?

__________ J

19.2

A large pot is placed on a stove and 1.2 kg of water at 14°C is added to the pot. The temperature of the water is raised evenly to 100°C just before it starts to boil. (a) What amount of heat is absorbed by the water in reaching 100°C? (b) The water then boils until all of it has evaporated, turning to water vapor at 100°C. How much heat does the water absorb in this process?

(a) __________ J
(b) __________ J

19.3 How much heat is required to change 1.83 kg of solid lead at 135°C to liquid lead at 327°C?

__________ J

19.4 Steam at 100°C is injected into 1.8 kg of water at 22°C in a well-insulated container, where it condenses and mixes with the existing water, reaching thermal equilibrium. If the final temperature of the well-mixed water is 25°C, what is the mass of the injected steam?

__________ kg

19.5 An aluminum container and the water in it are in thermal equilibrium at 28°C. The container has mass 0.13 kg and the water has mass 0.27 kg. 0.11 kg of ice at −13°C is added, and the system again reaches thermal equilibrium. (a) What is the final temperature of the system? (b) If any of the ice melts, what is the mass of the ice that melts? (Enter "0" if no ice melts.)

(a) __________ ° C
(b) __________ kg

19.6 A well-insulated container holds 1.50 kg of water at 22.0°C. A 2.98 kg copper block is heated in an oven, then completely submerged in the water. When the liquid water and the copper block reach thermal equilibrium, their common temperature is 42.0°C, but 0.0100 kg of the water has become water vapor at 100°C. What was the initial temperature of the copper block?

__________ ° C

Section 23 - Thermal conduction quantified

23.1 A window opening is 1.1 m by 1.3 m, and the wall thickness is 0.12 m. The temperature difference between the inside and outside is 28°C. What is the rate of heat transfer through the opening if it is filled with (a) concrete, or (b) wood (fir)?

(a) __________ W
(b) __________ W

23.2 A material has thermal conductivity of 0.380 W/m·K. What is its thermal resistance for a one-inch thick layer, in SI units of m^2·K/W?

__________ $m^2 \cdot K/W$

23.3 You want to add insulation under the roof of your house with a thermal resistance of 31.0 ft^2·F°·h/Btu. (a) What thickness of fiberglass should you add, in inches? (b) What thickness of plywood would do the job?

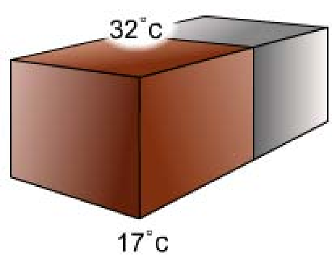
(a) __________ in
(b) __________ in

23.4 A circular pond of radius 12 m has a layer of ice 0.25 m thick. On a spring day, the temperature of the water below the ice is 4.0°C and the temperature of the air above is 11°C. What is the rate of heat transfer through the ice, in kW?

__________ kW

23.5 A cube of wood and a cube of concrete, each 0.17 m on a side, are placed side by side. One of the long faces of the rectangular prism formed by the two cubes is held at 17°C, and the opposite long face is held at 32°C. What is the total rate of heat transfer through the cubes?

__________ W

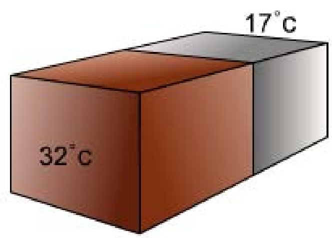


23.6 A copper rod and an aluminum rod with identical cross-sections are joined end to end. The free end of the copper rod is held at a constant temperature of 94°C, and the free end of the aluminum rod is held constantly at 13°C. Where the rods meet, the temperature is 45°C. If the copper rod is 2.4 m long, how long is the aluminum rod? (For the temperatures in the range discussed here, the thermal conductivity of copper is 401 W/m·K and the thermal conductivity of aluminum is 237 W/m·K. Ignore the heat loss through the sides of the rods.)

__________ m

Section 24 - Conduction through composite objects

24.1 A cube of wood and a cube of concrete, each 0.15 m on a side, are placed end to end. The far end of the wood cube is held at a constant 32°C and the opposite end of the concrete cube is at a constant 17°C. What is the rate of heat transfer through the pair of cubes?



__________ W

24.2 Two sheets of plywood, 1.2 by 2.4 m in dimension, are 0.013 m thick. They are separated by 0.076 m of fiberglass. If the plywood is replaced by two thinner sheets that are 0.0095 m thick, how thick should the fiberglass layer be to have the same thermal resistance overall?

__________ m

24.3 Three equally thick slabs of wood, glass and concrete are assembled in that order. The temperature on the outside surface of the wood layer is maintained at constant $44°\text{C}$, the outside surface of the concrete layer is a constant $16°\text{C}$, and the system is in steady-state (the internal temperatures are not changing). What is the temperature at the middle of the glass layer?

__________ ° C

24.4 A refrigerator has surface area of 6.96 m^2. The walls of the refrigerator are constructed of layers of polystyrene, urethane foam, and steel, from the inside out. The polystyrene is 0.0032 m thick, and has a thermal conductivity of 0.11 W/m·K. The urethane foam is 0.035 m thick, with a thermal conductivity of 0.060 W/m·K. The steel is 0.0018 m thick and has a thermal conductivity of 51 W/m·K. The inside temperature of the refrigerator is 2.4°C and the outside room temperature is 20.0°C. What is the amount of heat transferred through the walls of the refrigerator over a 24-hour day, assuming no one opens it?

__________ J

Section 27 - Radiation quantified

27.1 A spherical blackbody radiator has a temperature of 159°C. Its radius is 0.253 m. (a) At what rate does the sphere radiate energy? (b) If the sphere is in a small container whose walls are at temperature 19.0°C, what is the **net** rate of energy radiated from the sphere?

(a) __________ W
(b) __________ W

27.2 An object's temperature increases, and in so doing its radiation power increases by a factor of 2.00. What is the ratio of the new temperature to the original temperature (with both temperatures in kelvins)?

__________ to 1

27.3 A light bulb filament has an emissivity of 0.29. At a temperature of 2790 K, the filament radiates 55 W of power. What is the surface area of the filament?

__________ m^2

27.4 The power radiated by an object is 145 W. If the object were a perfect blackbody radiator (and otherwise was the same) it would radiate 567 W. What is the emissivity of the object?

____

27.5 As our bodies burn food as fuel, we radiate energy. Assume a person has a body surface area of 1.80 m^2 and an emissivity of 0.620 (a reasonable value for someone wearing clothes). What is the power radiated by the person? (Normal body temperature is 37.0°C.)

__________ W

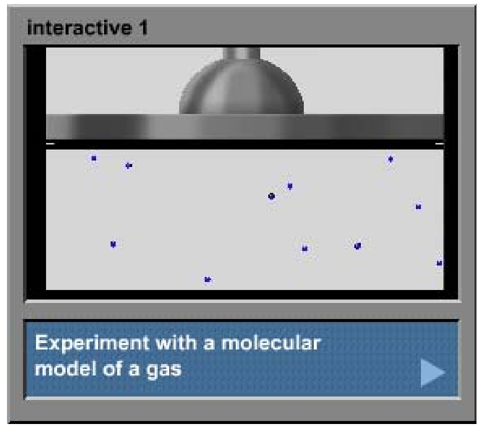
Section 29 - Interactive summary problem: pop the cork

29.1 Use the information given in the interactive problem in this section to calculate the heat flow required to pop the cork. Test your answer using the simulation.

__________ J

Chapter 20 Kinetic Theory of Gases

20.0 - Introduction



We live in a vast ocean of gas called the atmosphere. Deeper than the seas, it provides the oxygen we breathe, guards us against harmful radiation from space, helps to regulate the temperature, and makes flight possible.

In addition to supporting life, gases are used in many engines, and this is an important topic of study in physics. Transferring heat to a gas can cause it to expand. As the gas expands, it is able to do useful work. To understand how engines work, it is necessary to understand the behavior of gases.

In this chapter and elsewhere in this textbook, we focus on ideal gases. An ideal gas can be modeled as a number of small, hard spheres (atoms or molecules) moving rapidly around a container and colliding with each other and the container walls in perfectly elastic collisions. This aspect of the study of gases, which focuses on the motion of gas particles, is called the *kinetic theory of gases*.

To begin your study of gases, try the simulation to the right. In the container, each particle represents a gas atom or molecule. A gauge in the simulation provides a readout of the gas pressure inside the container. The pressure is a function of various factors including the volume of the container and the speed and the number of the particles.

The pressure measured in the simulation is caused by the collisions of the particles with the walls of the container. Each time a particle strikes a wall of the container, it exerts a force on the wall and the wall exerts an equal amount of force on the particle. The greater the speed with which the particle strikes the wall, the greater the force the particle exerts on the wall.

As detailed below, you can change three properties of the gas. The changes you make will be reflected solely in changes in the pressure of the gas. Other properties of the gas could also change, but in this simulation we hold those other properties constant.

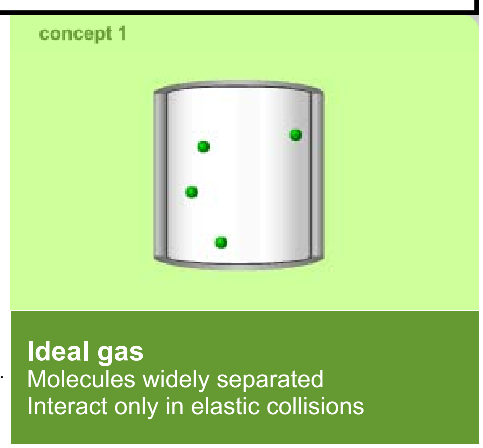
Clicking the up and down arrows for volume causes an external mechanism to raise or lower the container's lid. As the volume changes, the gas particles continue to move at the same average speed. How do you think gas pressure will relate to the volume of the container? Will decreasing the volume increase or decrease the pressure? Try it. Does the simulation help you see why there is a relationship between volume and pressure? Consider how it alters the frequency of the collisions.

You can also vary the temperature of the gas. The volume will not change as you do this, since the lid's position remains fixed. If you increase the temperature of the gas, you increase the average speed of the particles in the gas, and if you decrease the temperature, you decrease their average speed. Considering how the force and frequency of the collisions affects pressure, what do you think is the relationship between temperature and pressure when the volume is constant? Again, run the simulation to confirm your hypothesis.

The final controller allows you to add or subtract particles from the chamber. The particles you add will move at the same average speed as the gas particles already there. Do you think adding particles will result in more or less pressure? To test your conclusion, run the simulation to add particles and observe the resulting pressure.

When you have finished answering the questions above, you have experimented with some of the essential properties of a gas. The simulation should enable you to see relationships between gas pressure and the volume, temperature, and quantity of a gas.

20.1 - Ideal gas



Ideal gas: A gas that can be modeled as a set of particles having random elastic collisions. The collisions are the only interactions between the particles.

The concept of an ideal gas is used frequently in physics, both in the study of gases and later in the field of thermodynamics, where ideal gases play an important role. Real gases tend to behave like ideal gases at low enough densities (where quantum mechanical effects are not important). This makes ideal gases a good model for analyzing gas behavior. Here, we summarize the major properties defining an ideal gas.

Some gases, like carbon dioxide (CO_2) and oxygen (O_2), are composed of atoms bonded together in molecules. Other gases, like neon (Ne) and helium (He), are made up of particles that are individual atoms. For the sake of brevity, we will simply say that gases contain molecules (or particles), rather than repeatedly writing "atoms or molecules." When we analyze a gas, it is one made up of a single substance, such as oxygen or carbon dioxide.

An ideal gas consists of a large number of molecules. They are separated on average by distances that are large relative to the size of a molecule. They move at high speeds and collide frequently.

Using descriptions like "large" and "frequently" does not convey the nature of gases as well as specific numbers. Consider nitrogen molecules

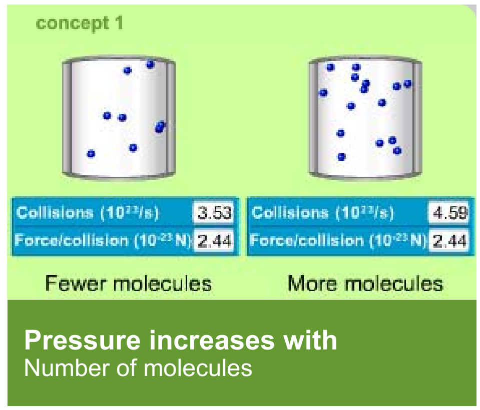
(N_2), the most common molecule in our atmosphere. In one cubic meter of nitrogen molecules at $0°C$ and one atmosphere of pressure, there are $2.69×10^{25}$ molecules. They occupy just 0.07% of the space and move at an average speed of 454 m/s. The molecules run into one another a lot: There are $9.9×10^{34}$ collisions each second between molecules in that space. The condition of $0°C$ (273 K) at one atmosphere is called *standard temperature and pressure* and is often used in calculations involving gases.

The molecules collide elastically with both each other and the container walls. The molecules only exert forces on one another when they collide. Forces that they exert on one another at other distances are negligible.

Each molecule can be considered as a small, hard sphere. In elastic collisions, both kinetic energy and momentum are conserved. The velocities of the molecules change after a collision, but the total momentum and kinetic energy of the molecules remain the same.

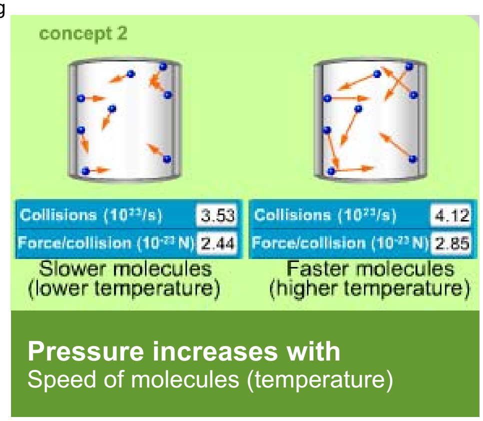
The walls of the container are rigid. When a molecule collides with a wall, its speed does not change although its velocity does (because its direction changes). Newtonian physics can be used to analyze the collisions of the molecules with each other and with the walls of the container.

20.2 - Gas pressure



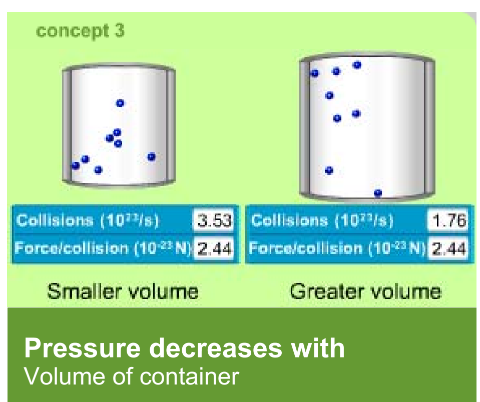
Gas molecules not only collide with each other, but also with the walls of any container. During these collisions, the molecules exert a force on the walls. Dividing the total amount of force caused by these collisions by the surface area of the walls yields the absolute pressure exerted by the gas.

In this section, we consider three factors that qualitatively influence gas pressure. If you tried the simulation in the introduction section, some of this discussion should be familiar to you. In this section, we show two containers of gas. For visual clarity, we show only a small fraction of the actual number of particles in the containers. Each particle in the simulation represents about $3.4×10^{18}$ molecules. We have done the calculations assuming that these are nitrogen molecules in a container about 10^{-2} m wide. In the container on the left in each diagram, the pressure is one atmosphere and the temperature is $0°C$. We vary one property at a time in each illustration.



Consider two containers of equal volume containing different numbers of gas particles moving at the same speed. This is shown in Concept 1. The pressure will be greater in the container on the right (the one that contains more molecules). Why? With everything else equal, the collisions are more frequent in the container with the greater number of molecules, which means the molecules collectively exert more outward force on the walls of the container. Greater average force on the same amount of surface area means the pressure is greater.

The relationship between molecular speed and pressure is shown in Concept 2. Now each container contains the same number of molecules, but the container on the right has faster moving molecules (at a greater temperature) and, thus, greater pressure. Why does pressure increase with the speed of the molecules? There are two reasons why this is the case. First, when molecules move faster, they collide more frequently with the walls of the container. Second, when a molecule strikes the wall and rebounds, the molecule exerts more force at greater speeds than at slower speeds. (Imagine the force involved in a tennis ball rebounding off you when it is thrown at 1 m/s versus 20 m/s.) Since the average speed of molecules increases with temperature, this means pressure also increases with temperature.



In Concept 3 the temperature and the number of molecules are the same for each container. We change the volume of the container on the right. The pressure **decreases** when the container volume **increases**; pressure is inversely proportional to volume. In a larger container, the molecules collide less frequently with the walls. The surface area of the container also increases, so pressure (force divided by area) decreases for that reason, too.

20.3 - Boyle's and Charles' gas laws

In this section we take a macroscopic view of ideal gases to discuss their overall properties. Several laws state the relationship among macroscopic properties such as pressure, volume, temperature, and the amount of gas (number of molecules). We start here with Boyle's and Charles' laws.

Boyle's law is named for Irish scientist Robert Boyle, who in 1662 discovered that the absolute pressure of a fixed amount of gas **at a constant temperature** is inversely proportional to its volume. To put it another way, if you decrease the volume of a gas, its pressure increases proportionally. Boyle's law is stated in Equation 1: The product of the initial pressure and volume of a gas equals the product of its final pressure and volume.

Jacques Charles, a Frenchman who worked a century later than Boyle, also studied gases. One way to state Charles' law is that for a fixed amount of gas **at a constant pressure**, the ratio of its volume to its absolute temperature remains constant. You see this in Equation 2 using the initial and final volumes and temperatures of a gas.

These gas laws are empirical; that is, they are determined to be true through experiment and measurement rather than being derived from other laws.

equation 1

Robert Boyle (1627 -1691)

Boyle's law

$$P_iV_i = P_fV_f$$

P = absolute pressure
V = volume
Temperature, quantity of gas constant

equation 2

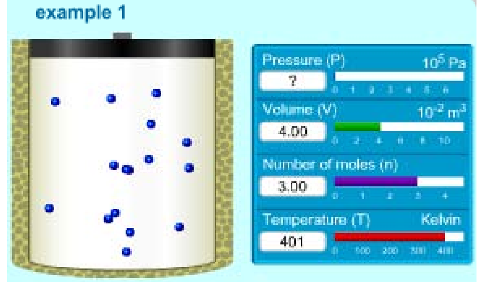
Jacques Charles (1746 - 1823)

Charles' law

$$\frac{V_i}{T_i} = \frac{V_f}{T_f}$$

V = volume
T = Kelvin temperature
Pressure, quantity of gas constant

example 1



0.020 m^3 of gas expands to 0.040 m^3. Its initial pressure was 5.0×10^5 Pa. Assume its temperature is constant. What is its final pressure?

$P_iV_i = P_fV_f$
$P_f = P_iV_i/V_f$
$P_f = (5.0\times10^5\text{ Pa})(0.020\text{ m}^3)/0.040\text{ m}^3$
$P_f = 2.5 \times 10^5\text{ Pa}$

20.4 - Avogadro's number and moles

Avogadro's number: 6.022×10^{23}.

Mole: The amount of a substance that contains Avogadro's number of particles.

The concept of moles is used in the study of gases. You might find that this section reviews what you learned in a chemistry class, where this concept also figures prominently.

Physicists often need to deal with large quantities of molecules when they analyze gases. For instance, one cubic meter of air at sea level contains about 3×10^{25} molecules. In addition, scientists may need to quantify a given amount of gas by its number of molecules rather than its mass. To do so, they use moles and Avogadro's number.

One mole is defined as the amount of a substance containing the same number of particles as there are atoms in 12 grams of carbon-12. Carbon-12 is one form, or isotope, of carbon, having six protons and six neutrons in its nucleus. Scientists have determined that there are, to six significant figures, $6.022\ 14\times10^{23}$ atoms in 12 grams of carbon-12, and have named this value Avogadro's number (after the 19th century Italian physicist Amedeo Avogadro). Avogadro's number is indicated by N_A.

A mole contains about 6.022×10^{23} particles of a substance. The abbreviation for mole is "mol". The concept of the mole is important enough that the mole is one of the seven fundamental units of the *Système Internationale*.

Avogadro's number specifies the number of particles per mole of a substance. It can be applied to any kind of object. For example, if you had Avogadro's number of golf balls, or one mole, you would have 6.022×10^{23} balls (and their mass would be 2.8×10^{22} kg).

Of course, moles are more typically used to specify the number of molecules in a quantity of gas. Moles are used to focus on the number of molecules, not their mass. For example, a mole of hydrogen atoms contains the same number of atoms as a mole of lead, even though one mole of lead atoms weighs 207 times more than the hydrogen atoms. If you have studied chemistry, you may be familiar with combining substances in terms of their molar amounts to achieve the correct ratio of molecules.

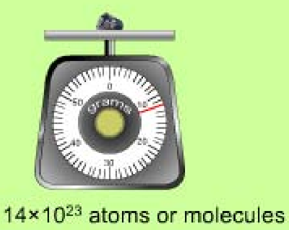
The following discussion will prove useful if you need to use atomic mass units. The *atomic mass scale* relates moles and the mass of individual molecules. The *atomic mass unit*, symbolized by u, is based on carbon-12. By definition, one atom of carbon-12 has a mass of exactly 12 atomic mass units, or 12 u. One atomic mass unit equals $1.660\ 54\times10^{-27}$ kg. Since Carbon-12 contains six protons and six neutrons, which have about the same mass and make up almost all the mass of the atom, the mass of a proton or neutron is approximately one atomic mass unit.

A particle with an atomic mass of 24 u is twice as massive as a carbon-12 atom. This means a mole of the 24 u substance will have twice as much mass as a mole of Carbon-12.

The mass per mole of a substance is called its *molar mass*. When measured in grams (or, to be precise, g/mol) the molar mass of a particular substance has the same numerical value as the atomic mass of the substance. Why? The atomic mass of an atom of carbon-12 is 12 u by definition, and a mole of carbon-12 atoms is exactly 12 grams, by the definition of Avogadro's number.

Consider a substance with an atomic mass of 36 u. Each particle will be three times as massive as an atom of carbon-12. One mole of the substance has the same number of particles as one mole of carbon-12. This means a mole of this substance would have a mass that is three times the mass of a mole of carbon-12, or 36 grams. In physics, the molar mass is usually stated in kg/mol instead of g/mol.

concept 1



Avogadro's number

Number of atoms in 12 g of carbon-12

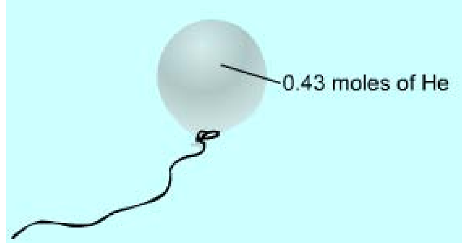
$N_A = 6.022\times10^{23}$ particles

concept 2

Mole

Contains Avogadro's number of particles

example 1



How many helium atoms are in the balloon?

(0.43 moles)(6.022×10^{23} atoms/mol)

2.6×10^{23} helium atoms

20.5 - Ideal gas law

Ideal gas law: The product of a gas's pressure and volume is proportional to the amount of gas and its temperature.

The ideal gas law shown in Equation 1 is an empirical law that describes the relationship of pressure, volume, amount of gas, and temperature in an ideal gas. It states that the pressure of an amount of an ideal gas times its volume is proportional to the amount of gas times its temperature. The pressure is the absolute, not gauge, pressure. (Absolute pressure equals gauge pressure plus atmospheric pressure.) The small n in the equation is the number of molecules, measured in moles. The temperature is measured in kelvins. R is a constant called the *universal gas constant*. Its value is 8.314 51 J/mol·K.

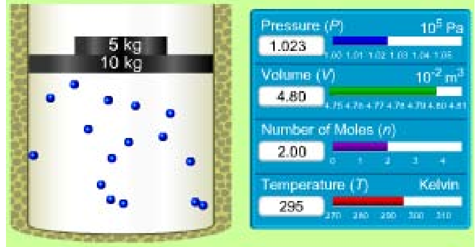
Everyday observations confirm the main precepts of the ideal gas law. When you inflate a bicycle tire, you increase the amount of gas in the tire, which increases both its pressure and volume. Once it is inflated and its volume is held nearly constant by the tire, an increase in temperature increases the pressure of the gas within. A hot day or the rolling friction from riding increases its temperature. You are told to check the pressure of a tire when it is "cold" since the correct pressure for a tire is established for 20°C or so. Checking tire pressure when the tire is hot may cause you to deflate the tire unnecessarily.

Sometimes, it is useful to be able to apply the ideal gas law when the number of particles N is known, rather than the number of moles. Since the number of particles is the number of moles times Avogadro's number, we can rewrite the ideal gas law as shown in Equation 2. The constant k here is called *Boltzmann's constant*, which equals the gas constant R divided by Avogadro's number N_A. Its value is $1.380\ 66\times10^{-23}$ J/K. While the first statement of the ideal gas law relates P, V and T to the **macro**scopic properties R and n of a gas, the second relates them to the **micro**scopic properties k and N.

If this is the ideal gas law, are there non-ideal gases? Yes, there are. For instance, water vapor (steam) is a non-ideal gas. The ideal gas law provides accurate results when gas molecules bounce into each other with perfectly elastic collisions. If the molecules also interact in other ways, as they do with steam, the law becomes a less accurate predictor of the relationship between pressure, volume, temperature, and the amount of gas. In the case of water vapor, *steam tables* are used to relate these values.

In Example 1, the volume of one mole of gas at standard pressure and temperature is calculated. The volume is 2.25×10^{-2} m^3 (22.5 liters, or about six gallons). One mole of any ideal gas occupies this same volume under these conditions. In 1811, Avogadro first proposed that equal volumes of all gases at the same pressure and temperature contained the same number of molecules.

concept 1



Ideal gas law

Pressure times volume proportional to:
· molar quantity of gas
· Kelvin temperature of gas

equation 1

Ideal gas law, number of moles

$$PV = nRT$$

P = absolute pressure (Pa)
V = volume, n = moles
R = universal gas constant
R = 8.31 J/mol·K
T = Kelvin temperature

equation 2

Ideal gas law, number of particles

$$PV = NkT$$

N = number of particles
$k = R/N_A$ = Boltzmann's constant
$k = 1.38\times10^{-23}$ J/K

example 1

What is the volume of one mole of gas at standard temperature and pressure? Standard temperature is 273 K and standard pressure is 1.01×10^5 Pa.

$$PV = nRT$$

$$V = \frac{nRT}{P}$$

$$V = \frac{(1.00\ \text{mol})(8.31\ \text{J/mol}\cdot\text{K})(273\ \text{K})}{1.01\times 10^5\ \text{Pa}}$$

$$V = 2.25\times10^{-2}\ \text{m}^3$$

20.6 - Sample problem: tank of air

The emergency scuba tank is filled with 8.50×10^{-2} m^3 of air at atmospheric pressure and 21.0°C. Immediately after the tank is filled, the absolute pressure is 1.53×10^7 Pa and the volume is 8.37×10^{-4} m^3. What is the temperature of the air in the tank at that time?

Scuba divers may carry small tanks like the one shown above to give them air in case of an emergency. Air that is at atmospheric pressure, 1.01×10^5 Pa, is pumped into the tank. The process of filling the tank increases the temperature and pressure of the air. The temperature will decrease as heat flows from the warmer tank to its cooler surroundings, but we assume the compression process occurs quickly enough for heat flow to be negligible. A value for atmospheric pressure is stated below.

Scuba diving equipment manufacturers use an estimated breath size of 1.6 liters, so this tank would hold a little over 50 typical breaths. Assuming a diver might breathe about 15 times a minute, an average under normal conditions, the emergency tank would give her around three minutes of air. However, she might be breathing a little more rapidly if she needs to use the tank!

Variables

initial volume of air	$V_i = 8.50\times10^{-2}\ \text{m}^3$
initial pressure	$P_i = 1.01\times10^5\ \text{Pa}$
initial temperature	$T_i = 21.0°\text{C}$
final volume of air	$V_f = 8.37\times10^{-4}\ \text{m}^3$
final pressure	$P_f = 1.53\times10^7\ \text{Pa}$
final temperature	T_f
moles of air	n

What is the strategy?

1. Model the air as an ideal gas. State the ideal gas law, and isolate the parameters that are constant, R and the amount n of gas, on one side of the equation.
2. Since the values mentioned above are constant, for both the initial and final states of the gas the "other" sides of the equations are equal to one another. Set them equal to each other in an equation.
3. Solve for the final temperature and evaluate.

Physics principles and equations

The ideal gas law

$$PV = nRT$$

The temperature is given in Celsius and will need to be converted to Kelvin to apply the ideal gas law.

Step-by-step solution

We start by using the ideal gas law to write an equation relating the initial and final conditions of the gas.

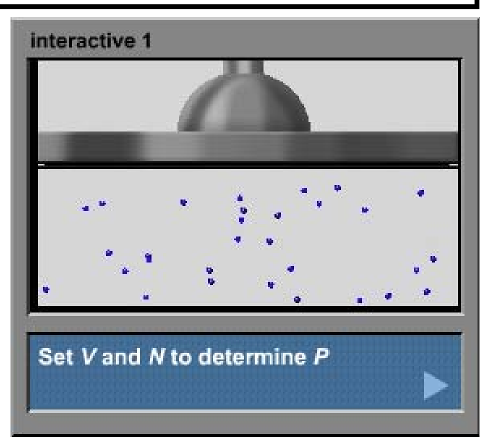
Step	Reason
1. $P_iV_i = nRT_i$	ideal gas law, initial condition
2. $\dfrac{P_iV_i}{T_i} = nR$	divide by temperature
3. $\dfrac{P_fV_f}{T_f} = nR$	ideal gas law, final condition
4. $\dfrac{P_iV_i}{T_i} = \dfrac{P_fV_f}{T_f}$	from steps 2 and 3

We can solve for the final temperature because all the other values are known. The initial temperature was stated in Celsius, but the ideal gas law requires that kelvins be used, so we add 273 to the Celsius temperature to convert.

Step	Reason
5. $T_f = \dfrac{P_fV_fT_i}{P_iV_i}$	solve for T_f
6. $T_f = \dfrac{(1.53\times10^{7}\ \text{Pa})(8.37\times10^{-4}\ \text{m}^3)(21.0+273\ \text{K})}{(1.01\times10^{5}\ \text{Pa})(8.50\times10^{-2}\ \text{m}^3)}$	substitute values
7. $T_f = 439$ K	evaluate
8. $T_f = (439 - 273) = 166°\text{C}$	convert to Celsius

This is very hot! When tanks like this are filled, they are usually held under cold water to increase the heat flow out of the tank and reduce their temperature increase.

20.7 - Interactive problem: pressure of an ideal gas



In this simulation, a small chamber contains an ideal gas consisting of 26 gas particles held at a constant temperature of 400 K. The gas has a pressure of 897 Pa. The chamber volume is currently set at 1.6×10^{-22} m^3.

You can change the volume of the chamber, within a range of 0.5×10^{-22} m^3 to 2.5×10^{-22} m^3. You also can vary the number of gas particles, from 10 to 30. Your task is to set the number of particles and the volume of gas so that the gas pressure is reduced to 552 Pa. The temperature of the gas in this simulation remains fixed at 400 K.

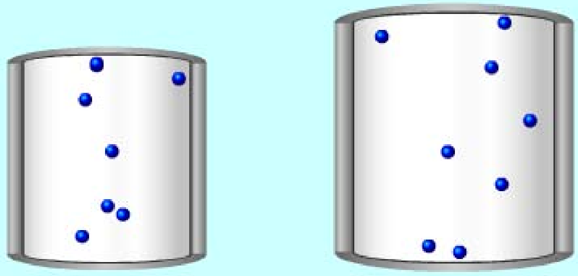
When you launch the simulation, you will see the gas particles moving around as determined by the settings stated above. After running for a few moments, the simulation will pause so that you can enter your values. Determine how to alter the volume and the number of molecules to achieve the target pressure of 552 Pa. Specify these values using the up and down arrows in the control panel and press GO to apply the new settings to the gas.

A text message will tell you whether the gas reached the correct pressure. You can also note the pressure in the pressure gauge in the simulation.

If you get the wrong answer, you can try again. Press RESET to return the gas to its initial conditions, specify new values, and press GO again.

If you have trouble with this simulation, review the section on the ideal gas law.

20.8 - Interactive checkpoint: volume of an ideal gas

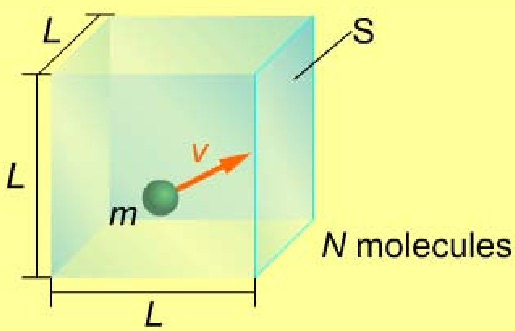


A quantity of an ideal gas occupies a volume of 2.18 m^3 at 288 K, with an absolute pressure of 1.45×10^5 Pa. If its temperature increases to 298 K and its pressure increases to 2.15×10^5 Pa, what is the volume?

Answer:

$V_f =$ ______ m^3

20.9 - Gas pressure, volume and kinetic energy



equation 1

Gas pressure, volume, and *KE*

$$PV = \frac{N}{3} m\overline{v^2} = \frac{2N}{3}\overline{KE}$$

P = gas pressure, V = volume
N = number of molecules
m = mass of one molecule
$\overline{v^2}$ = average squared molecular speed
$\overline{KE}$ = average kinetic energy

In this section, we demonstrate the relationship between two macroscopic properties of a gas − pressure and volume − and one of its microscopic properties, the average kinetic energy of its particles.

Earlier, we stated that the speeds of a gas's molecules are related to its temperature. In this section, we discuss the translational kinetic energy of the molecules, rather than their speed. Using Newtonian mechanics, we show a relationship between a gas's pressure and volume, and the kinetic energy of its molecules. If you like, you can view this as a derivation of a form of the ideal gas law, demonstrating a relationship between the pressure and volume of a quantity of gas and the energy of its molecules (which is a proxy for the gas's temperature).

For simplicity, we consider a container that is a cube. When we calculate the force exerted by one wall S of the cube on the gas, we will know the amount of force exerted by the gas on that wall, since Newton's third law states that these two forces must be equal and opposite. We can then divide by the surface area to calculate the pressure on the wall, and the result will be the equation stated above.

We will determine the force exerted on S by all the molecules in the cube by calculating the average force over time exerted by one molecule and then adding these individual forces for all the molecules. Because there are so many molecules, the overall force is constant, and we are justified in basing the calculation on an **average** force over time for a single molecule.

How do we determine the average force? Since the cube is filled with an ideal gas, the collisions of its particles with S are elastic. The force on the wall is a result of the momentum transferred when gas molecules collide with the wall. Each collision between a molecule and the wall results in a change in momentum for the molecule. We can use the concept of impulse to calculate the amount of force required to change the momentum.

Impulse is change in momentum, and also the product of the average force during the collision and the duration of the collision. This means the average force during a collision can be calculated as the change in momentum divided by the duration. In this case, though, we do not want to know the average force during the brief time interval of the collision. Rather, we wish to know the average force exerted by the molecule for the amount of time from one collision to the next.

In the derivation, we divide the change in momentum for one collision by the time between collisions, which is determined by how frequently a molecule collides with a wall of the container. That is the time Δt you see starting in step 3 below. This enables us to derive the average force over time exerted on wall S by one molecule.

Variables

length of edge of cube	L
velocity of a molecule	$\mathbf{v}$
mass of a molecule	m
number of molecules	N
momentum of a molecule	p
time between collisions	Δt
average force on S from one molecule	F_1
force on S from all molecules	F_S
volume of cube	$V = L^3$

Strategy

1. Determine the change in momentum when one molecule collides with wall S.
2. Determine the length of time between collisions of the molecule with the wall. From that, determine the average force over time of one molecule on the wall, as a function of the x component of the velocity of the molecule.
3. Calculate the average force on wall S from all the molecules.
4. Restate the force on wall S using the overall speed of the molecules, not just the x component of their velocities.
5. Use the definition of pressure to derive the desired equation.

Physics principles and equations

Definition of momentum

$$\mathbf{p} = m\mathbf{v}$$

We use the following equation relating the average force exerted on a molecule and its change in momentum.

$$\overline{\mathbf{F}} = \Delta\mathbf{p}/\Delta t$$

Definition of pressure

$$P = F / A$$

Mathematics principles

The overall speed of a molecule can be calculated from its components using the Pythagorean theorem. (The theorem is applied twice, first to the right triangle with legs v_x and v_y, and then to a right triangle where one leg is v_z and the other is the hypotenuse of the first triangle.)

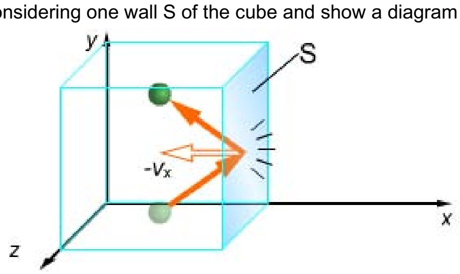
$$v^2 = v_x^{\,2} + v_y^{\,2} + v_z^{\,2}$$

Since the three components of the speed are perpendicular and independent, we can extend this relationship to average squared speeds.

$$\overline{v^2} = \overline{v_x^2} + \overline{v_y^2} + \overline{v_z^2}$$

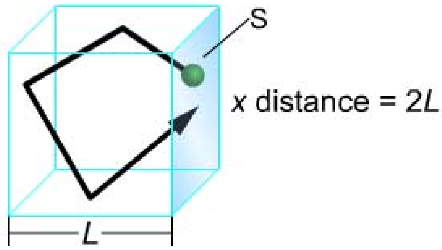
Step-by-step derivation

We start by considering one wall S of the cube and show a diagram of one molecule striking S.



When the molecule collides with the wall, it reverses the x component of its velocity vector. The other velocity components are not changed by the collision. From this, we can calculate the change in momentum of the molecule resulting from one collision with S.

We also want to know how much time elapses between collisions of the molecule with wall S. The following diagram shows that the distance in the x direction that the molecule travels between collisions is $2L$. For simplicity, we assume it does not collide with other molecules during its travels. The molecule may, however, bounce off other walls, which will not change its x velocity. (Motion in each dimension is independent of motion in the other dimensions.)



From this, we calculate the time between collisions. Dividing the change in momentum by this time value averages the force from one molecule on wall S. This is not the force during the collision, because we want to average the force over the entire time between collisions.

Step	Reason
1. $\Delta p_x = -mv_x - mv_x = -2mv_x$	change in momentum of molecule in x direction
2. $\Delta t = 2L/v_x$	definition of velocity
3. $F_1 = -\Delta p/\Delta t$	impulse equation
4. $F_1 = -(-2mv_x)/(2L/v_x)$ $F_1 = mv_x^2/L$	substitute equations 1 and 2 into equation 3 and simplify

Now we calculate the total force on wall S from all the molecules.

Step	Reason
5. $F_S = \sum_{i=1}^{N} \frac{mv_{ix}^2}{L}$	force from all molecules is sum
6. $F_S = \frac{m}{L} \sum_{i=1}^{N} v_{ix}^2$	factor
7. $\overline{v_x^2} = \left(\sum_{i=1}^{N} v_{ix}^2 \right) / N$	average of squared speeds
8. $\sum_{i=1}^{N} v_{ix}^2 = N\overline{v_x^2}$	rearrange
9. $F_S = \frac{Nm}{L} \overline{v_x^2}$	substitute equation 8 into equation 6

The previous equation is stated in terms of the x component of velocity. The motion of the molecules is essentially random, so **on the average** a molecule's speed is the same in all three directions. This gives us a way to restate the force using the overall molecular speed.

Step	Reason
10. $\overline{v^2} = \overline{v_x^2} + \overline{v_y^2} + \overline{v_z^2}$	Pythagorean theorem
11. $\overline{v_x^2} = \overline{v_y^2} = \overline{v_z^2}$	random motion of molecules
12. $\overline{v_x^2} = \overline{v^2}/3$	substitute equation 11 into equation 10
13. $F_S = \frac{Nm}{3L} \overline{v^2}$	substitute equation 12 into equation 9

Finally, we use the definition of pressure to find the pressure on wall S, and rewrite the resulting equation in the form we want to prove.

Step	Reason
14. $P = F/A$	definition of pressure
15. $P = F_S/L^2$	for wall S
16. $P = \frac{Nm}{3L^3}\overline{v^2}$	substitute equation 13 into equation 15
17. $P = \frac{Nm}{3V}\overline{v^2}$	volume of cube
18. $PV = \frac{N}{3}(m\overline{v^2})$	multiply by V

The right side of this equation contains as a factor the mass times the average squared speed of a molecule. This expression is twice the average kinetic energy of the molecule. This means the pressure times the volume is proportional to the average translational kinetic energy of a gas molecule, which is the relationship we wanted to demonstrate.

20.10 - Kinetic energy and temperature

In this chapter, we have shown that by varying the speed of its molecules, we can affect the pressure of a gas. We also have mentioned that the molecules' speed increases with temperature. Now we will get more specific about this relationship.

Rather than showing a relationship between speed and temperature, it proves more productive to show a relationship between translational molecular kinetic energy and temperature. To put it another way, we find it more useful to discuss the relationship between the molecular speed squared and temperature.

The molecules that compose a gas move at a variety of speeds. As the temperature of a gas increases, the average speed of its molecules increases, as does the molecules' average kinetic energy. The kinetic theory of gases links the temperature of a gas to the **average** kinetic energy of the molecules that make up the gas.

The relationship is quantified in Equation 1. The average translational kinetic energy of a particle in an ideal gas equals $(3/2)kT$, with k being Boltzmann's constant and T the Kelvin temperature. This equation results from combining the ideal gas law $PV = NkT$ with the equation derived in a prior section, $PV = \frac{2}{3}N\overline{KE}$.

The equation for the kinetic energy of a particle has an interesting implication: The kinetic energy of gas molecules is solely a function of temperature. It is independent of other factors, such as pressure or volume. As the graph in Equation 1 illustrates, when the temperature of a gas increases, the kinetic energy of the molecules that make it up increases linearly.

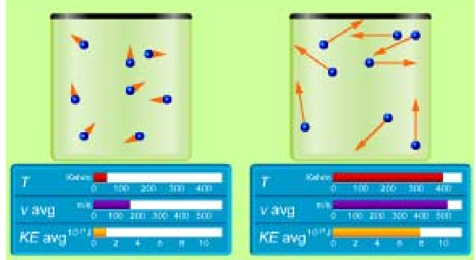
Now we move from considering the kinetic energy of a particle to the kinetic energy of an amount of an ideal monatomic gas. A monatomic gas is one whose particles are single atoms. The internal kinetic energy of a monatomic gas is the sum of the individual kinetic energies of all the atoms.

You see the equation for the internal energy of an ideal monatomic gas in Equation 2. It can be derived from the first equation by multiplying by the number of atoms. The kinetic energy of one molecule is $(3/2)kT$, as described above. For n moles of gas, there are nN_A atoms, so the total energy is $(nN_A)(3/2)\,kT$. Since k equals R/N_A, the internal energy of n moles of monatomic gas equals $(3/2)nRT$.

This energy makes up all the internal energy of an ideal monatomic gas, because the only form of energy of this type of gas is translational (linear) kinetic energy. There is no potential energy in any ideal gas. Since there are no forces between the molecules, there can be no energy due to position or configuration. The nature of monatomic molecules (single atoms) means they have neither rotational nor vibrational energy.

Although we have focused on ideal monatomic gases, the internal energy and temperature of ideal diatomic and polyatomic gases are also linearly related.

concept 1



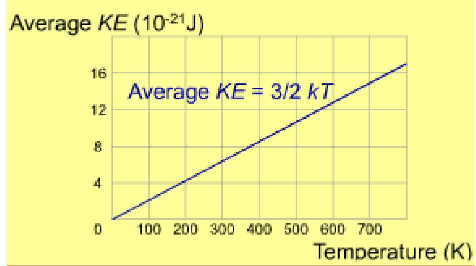
Kinetic energy of ideal gases

As temperature of gas increases:

· average molecular speed increases

· average translational *KE* of molecule increases

equation 1



Average *KE* of molecule, ideal gas

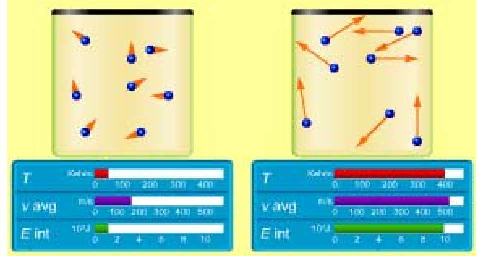
$$\overline{KE} = \frac{3}{2}kT$$

$\overline{KE}$ = average KE per molecule (translational)

k = Boltzmann's constant

T = Kelvin temperature

equation 2

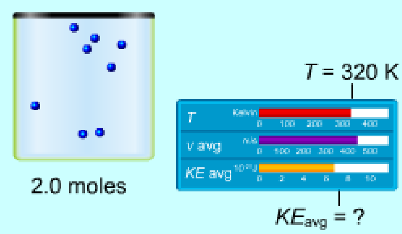


Internal energy, ideal monatomic gas

$$E_{\text{int}} = \frac{3}{2}nRT$$

E_{int} = internal energy of gas
n = amount of gas in moles
R = gas constant

example 1



Two moles of an ideal monatomic gas are at 320 K. What is the average *KE* of an atom?

$$\overline{KE} = \frac{3}{2}kT$$

$$\overline{KE} = \frac{3}{2}(1.38 \times 10^{-23}\ \text{J/K})(320\ \text{K})$$

$$\overline{KE} = 6.6 \times 10^{-21}\ \text{J}$$

example 2

What is the total internal energy of the gas?

$$E_{\text{int}} = \frac{3}{2}nRT$$

$$E_{\text{int}} = \frac{3}{2}(2.0\ \text{mol})(8.31)(320\ \text{K})$$

$$E_{\text{int}} = 8.0 \times 10^{3}\ \text{J}$$

20.11 - Root mean square

Root mean square: The square root of the average of the squares of a set of values.

The root mean square (rms) is used in several physics topics, including the study of gases. To calculate the root mean square of a set of numbers, square each of the numbers, calculate the average of those squares, and then evaluate the square root of that average. This equation is shown both in Equation 1 and in the example problem to the right.

With gases, the rms speed is used instead of the average speed in several equations. This is because we are interested in the average values of quantities, such as KE, that depend on the square of a property.

The average KE of 10 molecules could be calculated by calculating the KE of each molecule, and averaging those values. It could also be calculated by determining the rms speed of the molecules, and using that value for the speed in the expression $\frac{1}{2}mv^2$.

The example problem on the right asks you to calculate the rms speed for three molecules. The rms speed is 540 m/s, while the average speed is 539 m/s.

concept 1

Root-mean-square speed
Measure of speed useful in gas theory

equation 1

Root mean square

$$\text{rms} = \sqrt{\frac{\sum x_i^2}{N}}$$

rms = root mean square
x_i = one of a set of values
N = number of values

example 1

What is the root-mean-square speed of the three molecules?

$$\text{rms} = \sqrt{\frac{\sum x_i^2}{N}}$$

$$\text{rms} = \sqrt{\frac{(504\ \text{m/s})^2 + (525\ \text{m/s})^2 + (588\ \text{m/s})^2}{3}}$$

$$\text{rms} = \sqrt{292{,}000}\ \text{m/s}$$

rms = 540 m/s

20.12 - Rms speed of gas molecules and temperature

The kinetic energy of gas molecules is proportional to the temperature of the gas. This means the speed of the molecules must be related to temperature as well. An equation that relates speed, temperature and molar mass is shown in Equation 1.

The relationship between speed and temperature is stated using the root-mean-square (rms) speed of the molecules, not the average speed. The factor M is the molar mass of the gas measured in kg/mol.

As the equation indicates, at a given temperature, less massive molecules (ones with a smaller molar mass) move faster. Equation 2 includes a table of molar masses and rms speeds of some of the atoms and molecules that make up the Earth's atmosphere. All the speeds are calculated at the standard temperature of 273 K. The speeds range from xenon, which "dawdles" along at 227 m/s (817 km/h), to hydrogen, which moves at the brisk clip of 1840 m/s (6620 km/h).

equation 1

Root-mean-square molecular speed

$$v_{rms} = \sqrt{\frac{3RT}{M}}$$

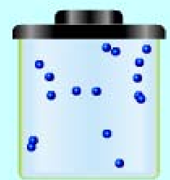
v_{rms} = root-mean-square speed
R = gas constant
T = Kelvin temperature
M = molar mass in kg/mol

equation 2

Gas *at 273 K*	Molar mass (kg/mol)	RMS speed (m/s)
Hydrogen (H_2)	0.00202	1840
Nitrogen (N_2)	0.0280	493
Oxygen (O_2)	0.0320	461
Argon (Ar)	0.0400	412
Carbon dioxide (CO_2)	0.0440	393
Ozone (O_3)	0.0480	377
Xenon (Xe)	0.131	227

Rms speeds for atmospheric gases

example 1



T = 303 K

The molar mass of nitrogen (N_2) is 0.0280 kg/mol. What is the root-mean-square speed of these nitrogen molecules?

$$v_{rms} = \sqrt{\frac{3RT}{M}}$$

$$v_{rms} = \sqrt{\frac{3(8.31\ \text{J/mol}\cdot\text{K})(303\ \text{K})}{(0.0280\ \text{kg/mol})}}$$

$$v_{rms} = 519\ \text{m/s}$$

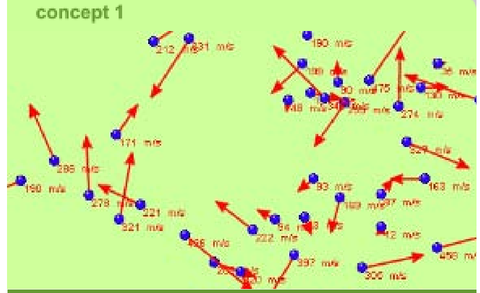
20.13 - Maxwell's speed distribution curve

Throughout this chapter, we have relied on the average or rms speed of molecules in a gas. This has been very useful and it provides essential insights. Here, we examine how the speeds of the molecules are distributed.

To explain the distribution of speeds in an ideal gas, we use the mathematics of probability and a distribution curve. Why? First, even small amounts of a gas are composed of an enormous number of molecules. For instance, a liter of the Earth's atmosphere contains about 3×10^{22} molecules. Second, the speed of any molecule in a gas changes frequently. In Earth's atmosphere, each individual molecule has over five billion collisions per second(!), leading to frequent changes in speed (and direction as well). These are two reasons why considering the distribution curve of the speeds of the molecules is appropriate and profitable.

The graph in Concept 2 represents the distribution of speeds in an ideal gas, in this case, oxygen at 300 K. This curve is called a *Maxwell distribution curve* after Scottish physicist James Clerk Maxwell, who pioneered research in this area.

concept 1



Maxwell speed distribution
Random collisions cause speeds of

The horizontal axis displays the possible speeds of molecules in meters per second. The vertical axis represents probability. The curved line is a probability distribution function, $y = P(v)$, and it is interpreted as follows. The area under any part of the curve bounded by two speeds represents the probability that the speed of a randomly chosen molecule will fall within that range of speeds.

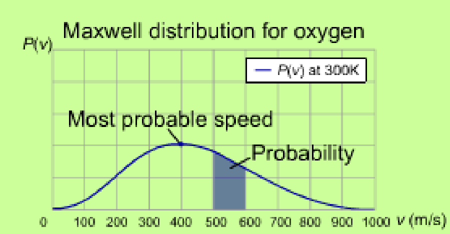
Since the area under the curve from 100 to 200 m/s is much greater than from the area under the curve from 0 to 100 m/s, the likelihood is much greater that at any instant in time, a molecule is moving at a speed between 100 and 200 m/s than between 0 and 100 m/s. In other words, the total number of molecules traveling at speeds from 100 to 200 m/s is greater than the total number of molecules traveling between 0 and 100 m/s.

The curve peaks at the most probable speed for a particle, but this speed is **not** the average speed of the molecules. That lies a little to the right of the most probable speed because the long "tail" of the curve on the right increases the average speed of the molecules.

The graph in Concept 3 compares the distribution of oxygen molecule speeds at two different temperatures. On average, the molecules at the higher temperature move faster. However, some molecules in the warmer gas move more slowly than some molecules in the cooler gas.

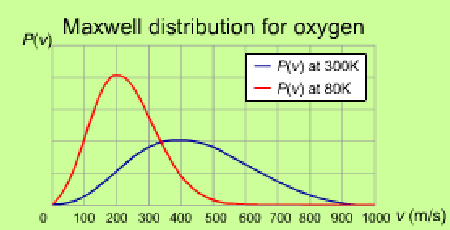
molecules to vary

concept 2



Maxwell speed distribution graph
Shows distribution of molecules' speeds
Represents probability molecule has speed in a given range
Peak is most probable speed

concept 3



Speed and temperature
Cooler gas: average speed decreases

20.14 - Interactive problem: speed distribution

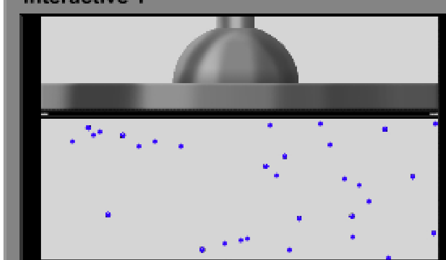
This simulation demonstrates how a Maxwell distribution curve accurately describes the distribution of particle speeds in an ideal gas.

The simulation consists of a small chamber of fixed volume that contains 50 gas particles. When you press GO, each of the 50 particles will begin moving at the same speed, 240 m/s, but in random directions.

A dynamic bar chart will graph the speeds of the particles. The chart displays ranges of particle speed on the horizontal axis (at the top), and it displays the percentage of particles traveling in each range as the height of a vertical bar. The simulation records the speeds of the particles about 20 times each second, and displays the cumulative speed data recorded from the beginning of the simulation. When the simulation starts, 100% of the particles are traveling at 240 m/s, but this will change rapidly.

Watch what happens to the chart of particle speeds as the particles begin colliding with one other. After 15 or 20 seconds, you will see a chart that approximates a Maxwell distribution curve.

interactive 1



Observe a Maxwell speed distribution

20.15 - Gotchas

Using the wrong temperature scale. Make sure you use Kelvin, not Celsius (and certainly not Fahrenheit!), when applying the equations in this chapter.

20.16 - Summary

In this chapter, you studied ideal gases, ones where the particles (atoms or molecules) can be modeled as interacting with each other and the walls of their container only in elastic collisions.

Avogadro's number is 6.022×10^{23}. It is a dimensionless number used to specify a quantity of matter. A mole contains Avogadro's number of particles.

The pressure of a gas increases with the number of molecules of the gas and its temperature, and decreases as its volume increases. The ideal gas law expresses this relationship in an equation. The ideal gas law can be stated in terms of the number of moles of a gas, or the number of particles.

The average kinetic energy of a gas molecule and the internal energy of a given quantity of gas can be calculated from its absolute temperature.

The root-mean-square (rms) speed of a gas molecule is used in some equations, and can be calculated from the gas temperature and molar mass.

The mean free path of a molecule is the average distance it travels before colliding with another gas molecule.

Maxwell's speed distribution curve expresses the probability that a gas molecule will have a speed in a given range. The probability depends on the temperature and molar mass of the gas.

Equations

Boyle's law

$$P_\mathrm{i}V_\mathrm{i} = P_\mathrm{f}V_\mathrm{f}$$

Charles' law

$$\frac{V_\mathrm{i}}{T_\mathrm{i}} = \frac{V_\mathrm{f}}{T_\mathrm{f}}$$

Ideal gas law

$$PV = nRT = NkT$$

Pressure, volume and energy (speed)

$$PV = \frac{N}{3}m\overline{v^2}$$

Energy and temperature

$$\overline{KE} = \frac{3}{2}kT$$

$$E_\mathrm{int} = \frac{3}{2}nRT \quad \text{(ideal monatomic gas)}$$

Particle speed

$$v_\mathrm{rms} = \sqrt{\frac{3RT}{M}}$$

Chapter 20 Problems

Chapter Assumptions

Use the following values for constants:

$N_A = 6.02 \times 10^{23}$

$R = 8.31$ J/mol·K

$k = 1.38 \times 10^{-23}$ J/K.

In problems which require you to know the atomic weights of atoms or molecules, use the following:

12.0 u for a carbon atom (C)

4.00 u for a helium atom (He)

1.00 u for a hydrogen atom (H)

14.0 for a nitrogen atom (N)

20.2 u for a neon atom (Ne)

16.0 u for an oxygen atom (O)

44.0 u for a carbon dioxide molecule (CO_2)

18.0 u for a water molecule (H_2O)

Section Problems

Section 3 - Boyle's and Charles' gas laws

3.1 To prepare for a magic trick, a magician squeezes an inflated balloon into a small box of volume 0.0760 m^3. If the balloon initially had a volume of 0.130 m^3 at an absolute pressure of 1.01×10^5 Pa, and the temperature of the gas inside it does not change, what is the pressure in the balloon after it is squeezed into the box?

__________ Pa

3.2 A scuba tank contains 2.83×10^{-2} m^3 of compressed air at an absolute pressure of 2.07×10^7 Pa. If all this air is released with no change in temperature, so that it has a pressure of 1.01×10^5 Pa (standard atmospheric pressure), what volume does it occupy?

__________ m^3

3.3 A balloon is filled with air outside on a very cold day, at an absolute pressure of 1.01×10^5 Pa and temperature −12.0°C. The balloon has a volume of 0.133 m^3. It is then brought indoors and its temperature increases to 22.0°C at the same pressure. What is the balloon's volume indoors?

__________ m^3

3.4 An experimental air compressor is designed to work by cooling the air to reduce its volume, without any change in pressure. To reduce the volume of a quantity of air at 34.0°C to 87.0% of its original volume, what temperature does the air have to reach?

__________ K

3.5 An air bubble released from a diving bell has a volume of 5.60×10^{-3} m^3 at a depth of 78.0 m below the surface. Assume the temperature of the air remains constant as it rises. What is the volume of the bubble when it reaches the surface?

__________ m^3

Section 4 - Avogadro's number and moles

4.1 How many molecules are there in 0.017 moles of carbon dioxide gas, CO_2?

__________ molecules

4.2 The molar mass of silver is 108 g/mol. What is the mass in kilograms of the 2.3×10^{24} atoms of silver that make up the chalice in the illustration?

__________ kg

4.3 The atomic mass of carbon is 12.0 u and of oxygen is 16.0 u, rounded to 3 significant figures. (a) What is the molecular mass of a carbon dioxide molecule, CO_2, in atomic mass units? (b) What is the molar mass of carbon dioxide in kg/mol? (c) How many molecules are in 0.0123 kg of carbon dioxide?

(a) __________ u
(b) __________ kg/mol
(c) __________ molecules

4.4 The molecular mass of water is 18 u. The mass of a teaspoon of water is 4.9×10^{-3} kg. If one teaspoon of water were spread evenly over the surface of the Earth, what would be the surface number density, that is, the number of water molecules per square meter? (Assume the Earth is a smooth sphere with radius 6.4×10^{6} m.)

__________ molecules/m^2

Section 5 - Ideal gas law

5.1 An ideal gas occupies 0.14 m^3 at an absolute pressure of 1.2×10^{5} Pa and temperature 280 K. (a) How many moles of gas are there? (b) How many molecules?

(a) __________ moles
(b) __________ molecules

5.2 At 265 K, the pressure of 0.0453 mol of an ideal gas is 1.42×10^{5} Pa. What volume does the gas occupy?

__________ m^3

5.3 A balloon holds 0.0140 m^3 of an ideal gas at a temperature of 35.0°C. If the amount of gas is 0.680 mol, what is the pressure of the gas?

__________ Pa

5.4 At a pressure of 2.32×10^{5} Pa, 6.41×10^{25} molecules of an ideal gas occupy 0.752 m^3. What is the temperature of the gas?

__________ K

5.5 A room with dimensions 2.52 × 4.57 × 5.61 m is filled with air at pressure 1.01×10^{5} Pa. Air can move into or out of the room through a crack under the door. The pressure stays constant as the temperature increases from 18.0°C to 24.0°C. What is the change in the number of moles of air in the room? Treat the air as an ideal gas. (Be careful with the sign of your answer.)

__________ mol

5.6 The atmosphere on Mars is almost all carbon dioxide (CO_2), and has a pressure of about 6.9×10^{2} Pa, much less than that of Earth. At −22°C (a warm day on Mars), what is the mass of one cubic meter of the atmosphere? (Assume the atmosphere is completely CO_2.)

__________ kg

5.7 A quantity of an ideal gas is held in a cylinder with a movable piston. At a temperature of 245°C, the gas has a volume of 0.0158 m^3 and pressure 2.12×10^5 Pa. The piston then moves so the volume of the gas is 0.0287 m^3 and the pressure is 1.67×10^5 Pa. What is the resulting temperature of the gas, in kelvins?

__________ K

Section 7 - Interactive problem: pressure of an ideal gas

7.1 Using the simulation in the interactive problem in this section, if there are 20 particles in the chamber, what volume will result in a pressure of 552 Pa? Calculate your answer first, then check it with the simulation.

__________ m^3

Section 10 - Kinetic energy and temperature

10.1 If 8.1 mol of an ideal gas occupy 0.26 m^3 at a pressure of 1.4×10^5 Pa, what is the average translational kinetic energy of a molecule of the gas?

__________ J

10.2 An ideal monatomic gas has an internal energy of 1.10×10^4 J. If there are 2.91 moles of the gas, what is its temperature?

__________ K

10.3 The average translational kinetic energy of a molecule of a quantity of an ideal gas is 1.10e−20 J. What is the temperature of the gas?

__________ K

10.4 Helium, the second most abundant element in the universe, can exist as a monatomic gas. At a pressure of 1.1×10^5 Pa, a quantity of helium gas occupies a volume of 0.16 m^3. What is the internal energy of the gas? Treat helium as an ideal gas.

__________ J

Section 11 - Root mean square

11.1 What is the root mean square of 29.2, 34.4, 18.6, 43.0 and 23.8?

____

11.2 The root mean square of five numbers is 27.4. Four of the numbers are 14.3, 33.7, 13.8, and 45.7. What is the fifth number?

____

Section 12 - Rms speed of gas molecules and temperature

12.1 What is the rms speed for a molecule of ozone gas (O_3) at standard temperature (273 K)?

__________ m/s

12.2 If the rms speed of the molecules of a quantity of helium at temperature 142 K is the same as the rms speed of neon atoms at a second temperature, what is the temperature of the neon gas?

__________ K

12.3 Two identical containers hold the same type of gas, at the same pressure and volume. The rms molecular speed for the gas in the first container is 637 m/s. The second container has twice as many molecules as the first. What is the rms molecular speed for the second container?

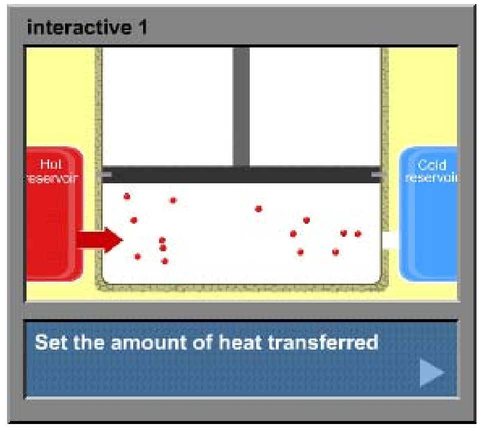
__________ m/s

12.4 On a particular afternoon in Houston, the nitrogen (N_2) molecules in the air have an rms speed of 509 m/s. What is the rms speed of the carbon dioxide (CO_2) molecules?

__________ m/s

Chapter 21 First Law of Thermodynamics, Gases, and Engines

21.0 - Introduction



The desire to build more powerful and efficient engines led engineers and scientists to embark on pioneering research into the relationship of heat, energy and work. The engines that powered the Industrial Revolution primarily used energy provided by burning coal.

Much has changed since the first engines were used to drain marshes and pump water from the coalmines of Great Britain. New sources of energy, from gasoline for automobiles to nuclear fuel for electric power generators, are used to power engines. Engines are now used in applications undreamt of by their first designers, for whom the horseless carriage and nuclear fission would have been science-fictional fantasies.

Despite the changes, much of the fundamental science developed in the 19th century is still used to analyze engines. The goal of an engine, to get useful work out of a heat source, has not changed. Although the technology has advanced and made engines more efficient, the physics of heat engines as established by 19th century engineers and scientists applies equally well to the steam locomotives of their era and to modern electrical power plants powered by nuclear fission.

This chapter focuses on two topics, using the engine as the basis of much of the discussion. One topic is the first law of thermodynamics, the relationship between the energy supplied to an engine and how much work it does. The other topic is the role of gases in the functioning of an engine. Many engines use a gas to function; applying some basic principles of how gases behave proves very useful in understanding the functioning of engines.

The simulation on the right will get you started on your study of engines. Here you see a heat engine, a device that uses heat as its energy source to do work. There is gas (represented by some bouncing molecules) in the central chamber of the engine, a piston on top, and a “hot reservoir” on the left from which heat can be allowed to transfer into the gas. In this case, we consider the cylinder, the gas and the piston to be the system. The reservoirs on the sides allow heat to flow in and out of the system. When heat flows into the engine’s gas, its temperature will rise (reflecting an increase in its internal energy), and if the piston is free to move, the gas will expand, pushing it up.

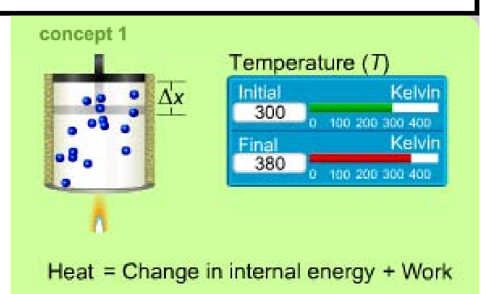
In this simulation, you will see what occurs during two distinct engine processes. In the first process, the piston is locked in place while heat is transferred to the gas. You control the amount of heat that is transferred. The internal energy of the gas is displayed in an output gauge. In the first step, you should ask yourself: What is the relationship between the amount of heat transferred to the gas and the change in the internal energy of the gas? Consider the principle of conservation of energy when you ponder your response to this question, and then test your hypothesis.

In the second process, no heat is allowed to flow in or out of the engine, but when you press GO, the piston will be unlocked so it can move. A gauge will show you how much work the gas does as it expands or contracts, changing the piston’s position. At the end of this process, note how much work the gas did and again its change in internal energy. How do the values for the work and change in internal energy relate during this process?

One last calculation (there are a few here, but you will have taught yourself the first law of thermodynamics when you complete this exercise): How does the amount of heat you initially transferred to the gas relate to its change in internal energy and the work it does? Write down these three values and see if you note a fundamental mathematical relationship. You can also apply the physics you learned earlier; consider the principle of conservation of energy and the work-energy theorem. (The work-energy theorem in its most general form says that the work done by an external force on a system equals its change in total energy, which includes thermal energy as well as mechanical energy.)

If these are too many questions, read on. The first law of thermodynamics begins this chapter.

21.1 - First law of thermodynamics



First law of thermodynamics
Heat transferred to system:
· increases internal energy and/or
· causes system to do work

First law of thermodynamics: The net heat transferred to a system equals the change in internal energy of the system plus the work done by the system.

The first law states that the net heat transferred into or out of a system equals the change in the internal energy of the system plus any work the system does. It is written as an equation in Equation 1, with Q representing the net heat transferred to the system. You saw this law at work in the simulation in the introduction to this chapter. First, heat was transferred to the gas in the engine, increasing the internal energy of the system. The increase in the temperature of the gas in the engine reflected the increase in internal energy. When you pressed GO again, the piston rose. The engine did work, and its internal energy decreased. The amount of work done by the engine equaled the magnitude of its change in internal energy as the piston rose.

We use the simple engine in Concept 1 to discuss the first law in more depth. Heat flows from the flame into the container, causing the internal energy of the gas it contains to increase. As the gas's internal energy increases, its temperature and the average speed of its molecules increases. As the gas's temperature increases, its pressure increases and it may expand. The gas does positive work when it expands, applying a force to the moving lid in the direction of its displacement.

It is important to define what we mean by "system." Here we define the system as the engine that includes the container, with its moveable lid, and the gas. When heat flows into the engine, the gas's temperature changes, but the container's temperature is assumed to stay the same. A rod is attached to the lid. The pressure of the gas pushes against the lid, and this can force both the rod and the lid upwards. The rod might be attached to the rest of the engine, such as the drive assembly of an automobile. When the gas expands, it does work on the rod.

The first law is mandated by the principle of conservation of energy. Conservation of energy applies to isolated systems, and the system of the engine is not isolated. But the engine is part of some larger system that is isolated. Heat energy flows into the engine. This energy must be conserved: Energy can be neither created nor destroyed in the process. That means the heat flow increases the internal energy of the gas. This energy, however, can do work, increasing the energy of some object outside of the engine. Work done by the gas reduces its internal energy.

The work done by the gas equals the force it exerts times how far the lid moves. That force equals the pressure times the surface area of the lid. Calculating the work done by the gas is easier when the pressure is constant than when the gas pressure varies.

Although it is not the typical way to analyze engines, you could think of the process entirely in terms of energy. Let's further simplify the engine, ignoring the rod and any forces acting on the lid other than gravity and the gas pressure from inside the engine. Heat flows in, raising the gas's internal energy and, if the lid rises, increasing the potential energy of the disk on top (by $mg\Delta y$). The sum of the increase in the gas's internal energy and the increase in potential energy of the disk equals the amount of net heat flow.

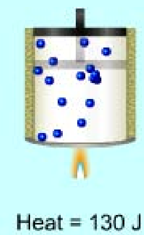
equation 1

First law of thermodynamics

$$Q = \Delta E_{int} + W$$

Q = net heat transferred to system
ΔE_{int} = change in internal energy
W = work done by system

example 1



The internal energy of the gas increases 75 J. How much work is done by the gas?

$Q = \Delta E_{int} + W$
$W = Q - \Delta E_{int}$
$W = 130\text{ J} - 75\text{ J}$
$W = 55\text{ J}$

The first law applies to any engine process, for any initial and final state of the gas plus any work done. For instance, in the simulation in the introduction, heat was added and then the piston was allowed to move, but this could have happened simultaneously, or it could have happened in a series of steps.

Mathematical signs are important here. A net flow of heat **into** the engine means Q is positive; a net flow of heat **out** of the engine means Q is negative. Work done by the gas raising the lid is positive, since the force is in the direction of the displacement. When the lid moves down, the work done by the gas is negative.

The relationship between heat, work and internal energy is illustrated in the example problem to the right. If 130 J of heat is transferred to an engine and its internal energy increases by 75 J, the first law of thermodynamics dictates that it must do 55 J of work. The change in internal energy and the work done by the gas must add up to equal the net heat transferred.

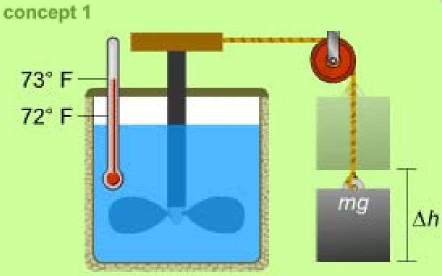
21.2 - James Joule and the first law

Although it is accepted today as one of the fundamental tenets of science, the proponents of the first law of thermodynamics faced a skeptical scientific community. Most scientists did not view heat as another form of energy. The prevalent theory in the early 1800s was the *caloric theory*, which proposed that a *caloric fluid* was added to matter as it was heated. This addition of fluid explained why a metal rod expanded when it was heated: It expanded because it contained more caloric fluid. One legacy of this belief is that heat still has its own units, the *calorie* and the *British thermal unit* (Btu).

Between 1843 and 1850, the English scientist James Joule performed a series of experiments that showed that heat was another form of energy. He used equipment similar to the apparatus shown in Concept 1. Rather than using heat to raise the temperature, Joule used mechanical energy. As the diagram illustrates, a falling weight causes the paddles to rotate, and the work of the paddles on the water causes its temperature to increase.

Joule applied the work-energy theorem to conclude that the work done by the paddles equaled the amount of change in the potential energy of the weight. By measuring the change in height and the mass of the weight, he could quantify the change in its PE and so the work done by the paddles. He also measured the increase in the temperature of the water.

concept 1



Heat as energy

Joule showed heat is a form of energy
Change of *PE* proportional to:
· water temperature increase
· heat required for same increase

Joule performed his experiment repeatedly. He showed beyond doubt that the change in the potential energy of the weight, which equaled the amount of work done by the paddles, was always proportional to the increase in the temperature of the water. This proved a relationship between mechanical energy and temperature.

Joule used British units. In his experiments, he calculated that 772.5 foot-pounds of work would increase the temperature of one pound of water by one degree Fahrenheit. Today, that value has been refined to 778 foot-pounds. Joule's measurement was impressively accurate.

Joule also knew how to use the specific heat of water to relate an amount of heat transferred to a quantity of water to its increase in temperature. He knew how much heat it would take to raise the temperature of a given amount of water by one degree, and with his apparatus, he could also calculate how much mechanical energy it would take to do the same. He showed that one calorie of heat equals an amount of mechanical energy that today we would describe as 4.19 J. (Since scientists today know heat is just one form of energy, they do not use the calorie unit as often, but instead use the joule as the unit for all forms of energy, as well as for work.)

Modern scientists still differentiate between heat (the transfer of energy) and the internal energy of an object. Heat changes the internal energy of an object or system, just as work does. Work and heat are two ways to change some form or forms of an object's energy. They reflect a process. It does not make sense to refer to "the work of an object" or "the heat of an object." Rather, it is correct to state how much work is done on or by an object, or how much heat is transferred into or out of the object. The result is a change in the object's energy.

21.3 - Heat engines

A heat engine uses the energy of heat to do work. Many engines have been designed to take advantage of heat energy. To cite two: A steam engine in an old locomotive and the internal combustion engine in a modern automobile both use heat as the source of energy for the work they do.

A heat engine is shown on the right. It is the container with a lid and piston on top, between a hot reservoir on the left and a cold reservoir on the right. In the engines we will consider, the container encloses a gas. The gas is cooler than the hot reservoir but warmer than the cold reservoir. Heat flows spontaneously from the hot reservoir into the gas in the engine, where the energy can be used to do work. Heat will also flow spontaneously from the gas to the cold reservoir. Otherwise, the container is insulated and no heat flows through any other mechanism into or out of the engine.

The reservoirs are large enough that they can supply or absorb as much heat as we like. We control the flow of heat between a reservoir and the container by opening a hole in the insulating wall to allow heat to flow, and closing it to stop the flow. The reservoirs are **not** part of the system we consider when applying the first law of thermodynamics.

The amount of gas in the engine stays constant. The temperature and pressure are assumed to be the same everywhere in the gas at any moment. When we analyze a heat engine, we will assume it is ideal: that there is no friction to reduce its efficiency and that heat flows uniformly and instantly through the system.

The engine goes through a sequence of processes that collectively are called an *engine cycle*. (One process or step in an engine cycle is sometimes called a *stroke*.) At the end of each cycle, the engine returns to its initial state. For instance, during one process, heat might flow into the engine from the hot reservoir so the piston rises. During another process, heat might flow out to the cold reservoir as the piston falls. No matter what the processes, at the end of an engine cycle, the engine returns to its initial configuration: The piston is in its initial position and the gas is back to its initial volume, pressure and temperature.

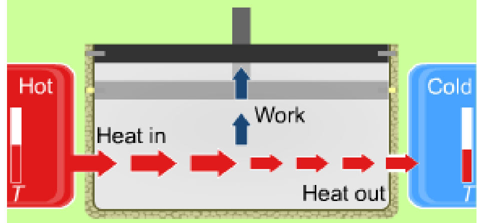
Since the gas returns to its initial state, its internal energy does not change during a complete engine cycle. This means we can concentrate on the two other quantities in the first law, namely heat and work. To apply the first law and other equations, we consider the work done **by** the gas. The gas does positive work as it lifts the piston. If the piston moves down and compresses the gas during another part of the cycle, the gas does negative work.

The distance the piston moves up or down might be the same. However, the amount of work that occurs will typically differ, since the purpose of the engine is to do net positive work on the piston during a complete engine cycle. The gas applies more force when it expands (to move a crankshaft, for example) than when it contracts, so the **net** work done by the engine in a cycle is positive. The net transfer of heat into the engine during a complete engine cycle equals the net work it does. The net transfer of heat equals the heat in from the hot reservoir minus the heat out to the cold reservoir.

You see this stated as an equation to the right. The net work done by the engine during a cycle equals the heat transferred into the engine from the hot reservoir minus the heat that flows out to the cold reservoir. Heat that flows into the engine from the hot reservoir is called Q_h, and heat that flows out to the cold reservoir is called Q_c. We treat both Q_h and Q_c as positive quantities. This means the **net** flow of heat equals $Q_h - Q_c$. If heat flows only out of the engine during a process in an engine cycle, then the net heat flow for that process is negative.

This equation is a special case of the first law of thermodynamics. The law states that the net heat flow equals the work done plus any change in internal energy. Since the internal energy is not changed after a complete engine cycle (the engine returns to its initial state), the net work done by the engine equals the net flow of heat.

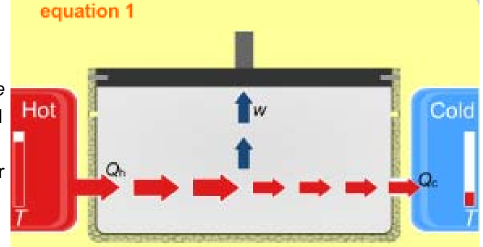
concept 1



Heat engines

Heat flows into engine
Engine uses heat to do work
Heat flows out of engine
Engine cycle: system returns to initial state

equation 1



For a complete engine cycle

$$W = Q_h - Q_c$$

W = net work done by engine
Q_h = heat in
Q_c = heat out
Internal energy returns to initial value

example 1

During an engine cycle, the heat transfers are as shown. What is

the net work done by the engine?

$W = Q_h - Q_c$

$W = 3200\text{ J} - 1800\text{ J}$

$W = 1400\text{ J}$

21.4 - The ideal gas law and heat engines

Many engines use a gas as their *working substance.* Heat is transferred to the working substance, which expands and does work. Water (both in liquid and gaseous form) is another common working substance. However, water vapor is far from an ideal gas, and in this textbook, we focus on an ideal gas as the working substance.

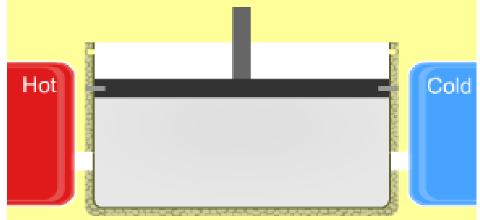
In addition to the first law of thermodynamics, the ideal gas law is very useful in analyzing the processes in an engine cycle. Here, we want to briefly review this law, and show how it is used in analyzing heat engines. The ideal gas law relates pressure, volume, temperature and the amount of gas. You see the law stated in Equation 1.

In the heat engines we consider, the quantity of gas enclosed by the container is constant. This leaves three variable properties of a gas in the ideal gas law: the gas's pressure, volume and temperature. The product of the pressure and volume is proportional to the temperature. If the temperature of the gas increases, for instance, then the product of its pressure and volume must increase, as well.

The first law of thermodynamics states that when heat flows into an engine, the heat energy increases the gas's internal energy and/or causes it to do work. A change in internal energy will be reflected in the gas's temperature; greater internal energy means a higher temperature. This means the volume or the pressure of the gas will increase, or both.

To apply these two principles, consider Example 1. The piston is locked in place so that the gas can do no work. This means that the heat transferred to the engine solely increases the internal energy of the gas. That increase in internal energy is reflected by an increase in temperature. Since the piston is locked, the volume of the gas is constant and the ideal gas law enables us to conclude that its pressure must increase proportionally to the temperature increase.

equation 1



Ideal gas law

$PV = nRT$

When n is constant:

$PV \propto T$

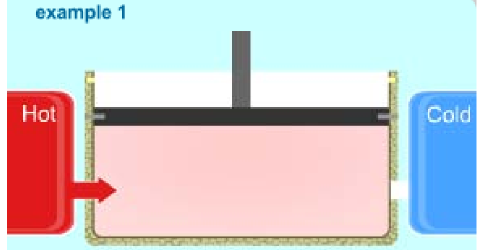
P = pressure, V = volume

n = number of moles of gas

R = gas constant

T = temperature (K)

example 1



The engine piston is locked in position and heat flows into the gas. What happens to the pressure, volume, and temperature?

Volume constant (piston locked)

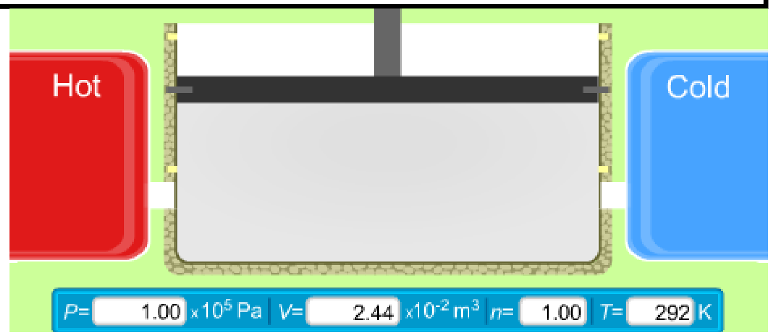
Temperature increases

Pressure increases proportionally

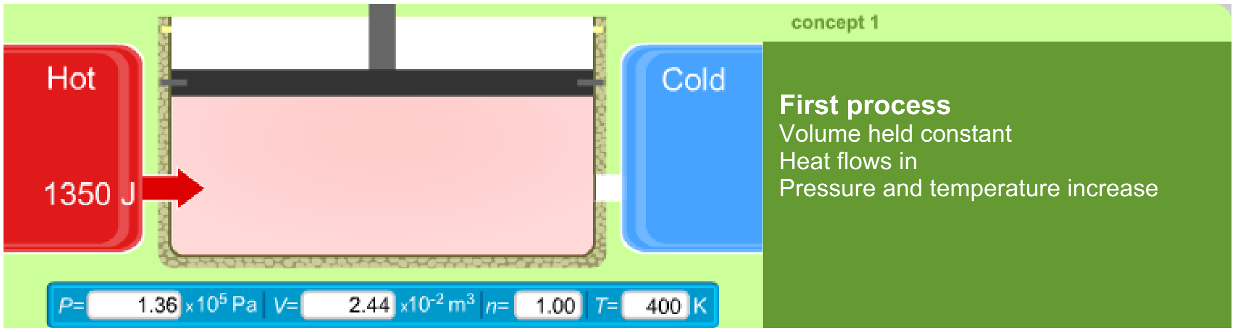
21.5 - Engine processes

In this section, we will review an engine cycle. It involves the three processes described below. To understand these processes, watch the gauges that display the pressure, volume, molar quantity and temperature of the gas. We assume that the pressure and temperature are uniform throughout the gas.

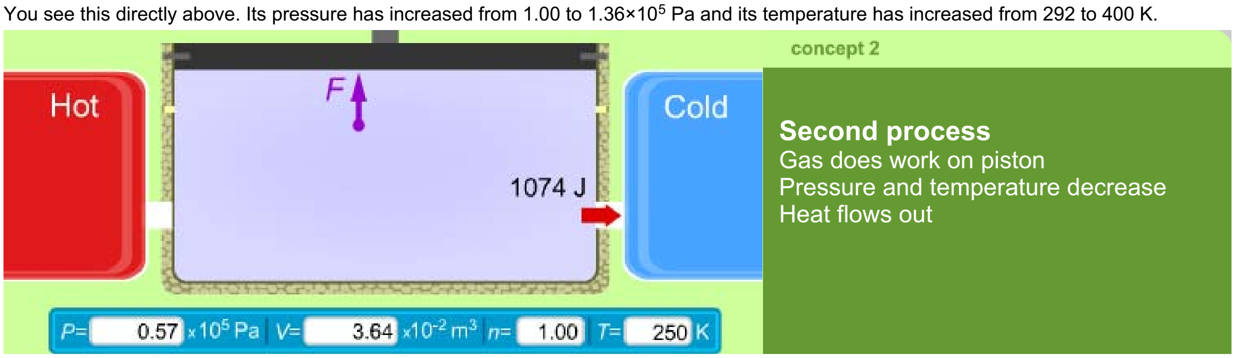
We start with one mole of gas with a volume of 2.44×10^{-2} m^3 at 292 K and a pressure of 1.00×10^5 Pa. The piston is locked and no heat is allowed to flow in or out. This state of the engine is depicted in the illustration at the top of the page.



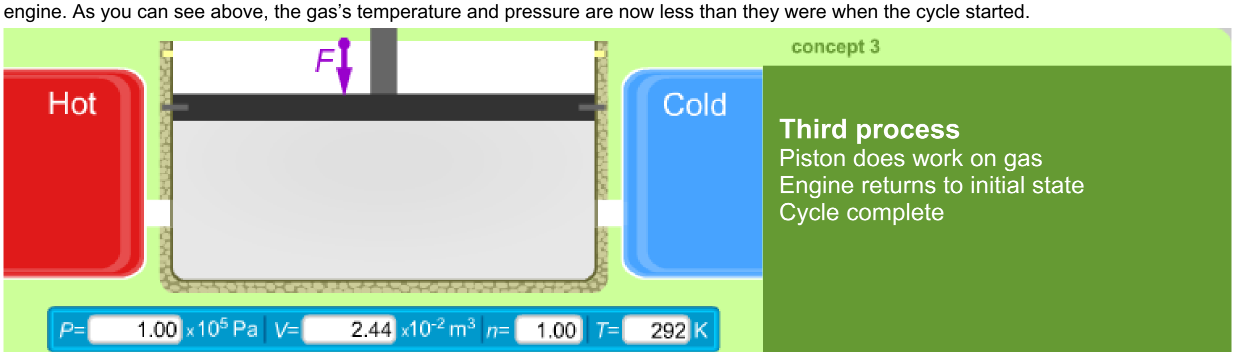
Initial state of heat engine.



First process. Now, 1350 joules of heat are allowed to flow into the gas. The energy inflow increases the internal energy of the gas, which is reflected by its increased temperature. Since the piston is locked, the gas's volume cannot change. This means its pressure must increase. You see this directly above. Its pressure has increased from 1.00 to 1.36×10^5 Pa and its temperature has increased from 292 to 400 K.



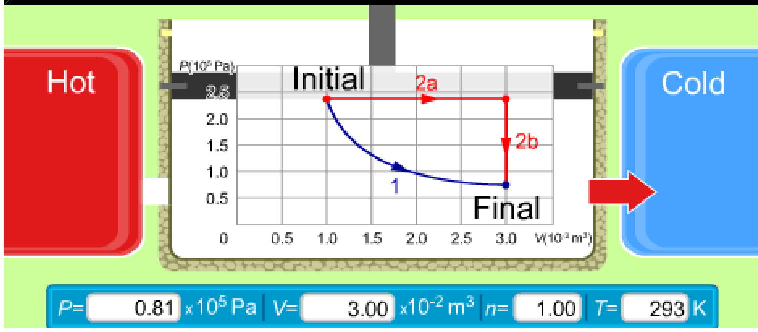
Second process. Now, we release the lock on the piston. This allows the gas to expand, doing work by moving the piston. The internal energy of the gas and its temperature decrease. A hole in the insulating wall opens to the cold reservoir, and 1074 J of heat also flow out of the engine. As you can see above, the gas's temperature and pressure are now less than they were when the cycle started.



Third process. The insulating walls to the reservoirs are again closed, so no heat flows. The piston pushes down, returning the gas to its initial volume. This is a lesser amount of work than was done when the piston moved up, because the gas pressure is always less during this process than during the expansion process. This work done on the gas increases its temperature to its initial value. The pressure returns to its initial state.

This is one of many possible engine cycles. We will examine some specific engine cycles in more detail later. For now, we focus on the thermodynamic processes that can make up part of an engine cycle.

21.6 - Pressure-volume graphs and heat engines

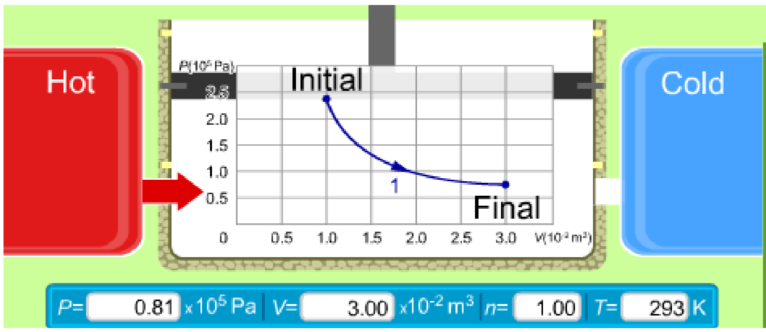


concept 1

Pressure-volume graphs
Multiple paths between same pressure-volume points

In an engine, the pressure and volume of the gas change as heat is transferred to the engine and the engine does work. These changes can be tracked with what is called a pressure-volume graph. In a pressure-volume graph, pressure is plotted on the vertical axis and volume on the horizontal axis. You see two engine processes diagramed above, one in blue and the other in red. They illustrate how two different processes can cause a gas to move between the same initial and final states.

We will now describe these processes step-by-step. The gas starts at 2.43×10^5 Pa of pressure, 1.00×10^{-2} cubic meters of volume, and a temperature of 293 K.

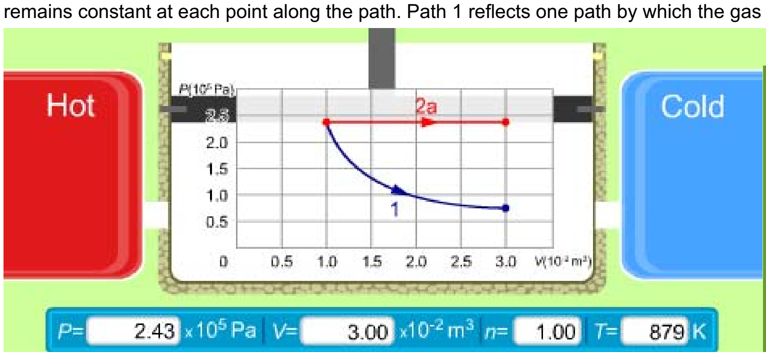


concept 2

Process 1
Volume increases as pressure decreases
Temperature does not change

Process 1. The gas's pressure is then reduced as its volume is increased in the process labeled 1, which is immediately above. The gauges and piston position in Concept 2 show the state of the engine at the end of the process. The temperature does not change during this process.

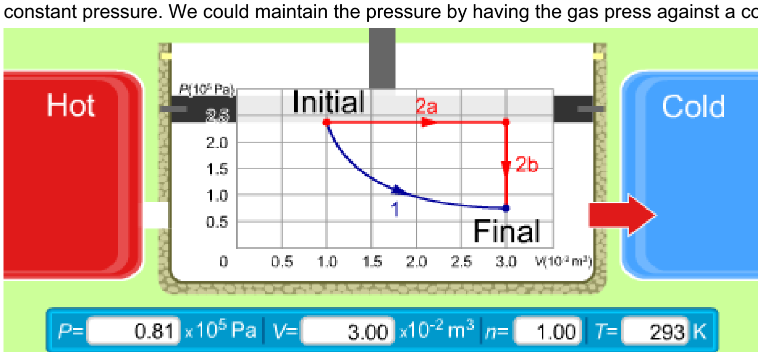
During the first process, the amount of heat added to the engine equals the work done on the piston because the internal energy and temperature of the gas do not change. The ideal gas law must be obeyed, so as the volume increases, the pressure decreases; their product remains constant at each point along the path. Path 1 reflects one path by which the gas can change from its initial to its final state.



concept 3

Process 2
a. Volume increases at constant pressure; temperature increases

Process 2, step a. Now we show a different process by which the gas can be caused to move between the same two states. To make the comparison clearer, we continue to show path 1 on the graph. The second process starts with the same pressure, volume and temperature as before. This process has two distinct steps. In the first step, labeled 2a in the graph above, heat energy flows into the gas from the hot reservoir as before. This increases the internal energy of the gas (as indicated by the increase in temperature). The graph is horizontal, indicating a constant pressure. We could maintain the pressure by having the gas press against a constant weight.



concept 4

Process 2
b. Pressure decreases at constant volume; temperature returns to initial

Process 2, step b. Next, since we have the same volume that was reached by the path 1 process, we lock the piston into place, keeping the volume constant. The pressure is too high, so we allow heat to flow out to the cold reservoir. This decreases the temperature of the gas, and at this fixed volume, this means the pressure must decrease proportionally. You see this step labeled 2b in the graph above.

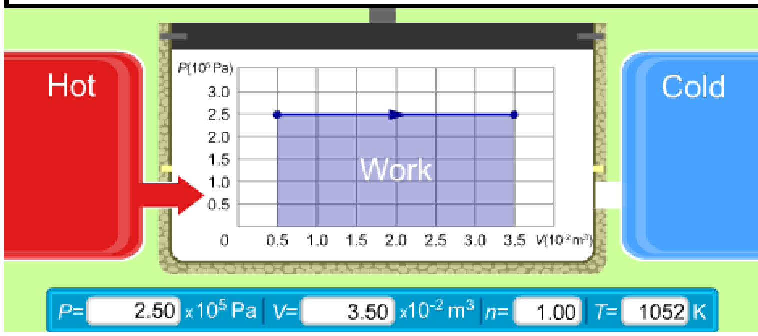
This two-step process arrives at the same final pressure-volume point as the first process, but in a different way. In the first, the gas's temperature remained constant while its pressure and volume constantly changed. In the second, first the volume was changed at constant pressure, and then the pressure was changed at constant volume. The temperature changed during both of the steps of process 2.

The difference in the paths reflects an important point: Gases can change from one combination of pressure and volume to another by experiencing different histories.

Both the processes sketched above obey the first law of thermodynamics and the ideal gas law. The first law is obeyed because there is the same **net** flow of heat into the engine in both cases. In process 1, all the heat is used for work, and the gas's internal energy stays the same. (To make the engine realistic, some heat should flow to the cold reservoir, since no engine is 100% efficient.) In process 2, more heat is added during step 2a than in process 1, since in this step the gas does the same amount of work as in all of process 1 **and** its internal energy increases. Heat flows out of the engine during step 2b, making the net heat into the engine the same for processes 1 and 2.

The ideal gas law is also obeyed. The product of the pressure and volume is always proportional to its temperature. For instance, on path 2b, as the gas's temperature decreases, its pressure decreases proportionally.

21.7 - Work and pressure-volume graphs



concept 1

Work and pressure-volume graphs

Amount of work = area under graph

Gas does

· Positive work as volume increases

· Negative work as volume decreases

The amount of work done by a gas can be analyzed graphically using a pressure-volume graph. The amount of work equals the area under the line or curve from the initial pressure-volume point to the final pressure-volume point. If the path of the process is from left to right, the volume increases and the gas pushes up the piston, doing positive work. If the path is from right to left, the volume decreases, the piston lowers, and the work done by the gas is negative.

We demonstrate the relationship between the area under the pressure-volume graph and work using the graph shown above. The pressure-volume path is a straight horizontal line from left to right. In this case, the volume increases (so the gas does positive work) but the pressure remains constant. Heat flowing into the engine allows it to do this work. We demonstrate that the work done is the area under this graph in the following steps.

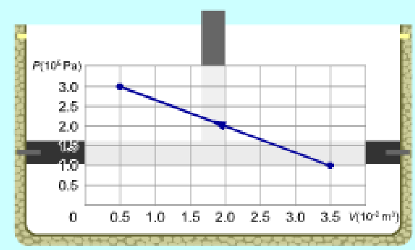
Variables

pressure of gas	P
change in volume of gas	ΔV
area of piston	A
force on piston	F
distance piston moves	Δy
work done by gas	W

Step-by-step derivation

Step	Reason
1. $P = F/A$	definition of pressure
2. $\Delta V = A\Delta y$	change in volume
3. $P\Delta V = (F/A)(A\Delta y) = F\Delta y$	multiply equation 1 by equation 2, simplify
4. $P\Delta V = W$	definition of work
5. $P\Delta V =$ area under graph	area of rectangle
6. $W =$ area under graph	substitute equation 4 into equation 5

example 1



Pressure increases from 1.00×10^5 to 3.00×10^5 Pa as the volume decreases from 3.50×10^{-2} to 0.50×10^{-2} m^3. How much work does the gas do?

$W = -$area under graph

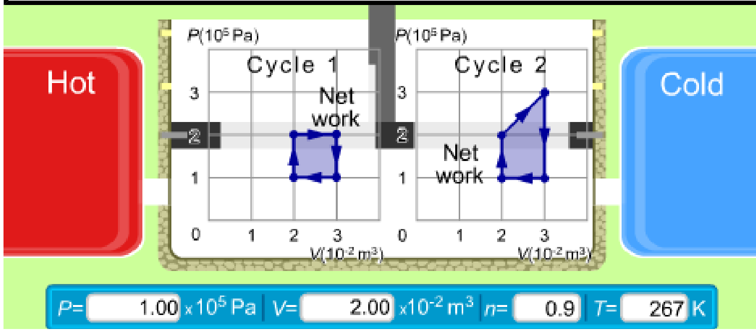
$A_{\blacksquare} = (3.00\times10^{-2}\ \text{m}^3)(1.00\times10^5\ \text{Pa})$

$A_{\blacktriangle} = \frac{1}{2}(3.00\times10^{-2}\ \text{m}^3)(2.00\times10^5\ \text{Pa})$

$W = -6.00\times10^3$ J

Pressure-volume graphs will not always be horizontal lines. In Example 1 to the right, you are asked to compute the work for a graph that is a straight line but not horizontal. The shape under this graph is a trapezoid. (You could also think of it, as we do in solving the problem, as a rectangle plus a triangle.) If we were measuring the work under a curve, we might have to approximate that area with a set of rectangles, or use calculus to calculate it.

21.8 - Pressure-volume graphs, engine cycles and work



concept 1

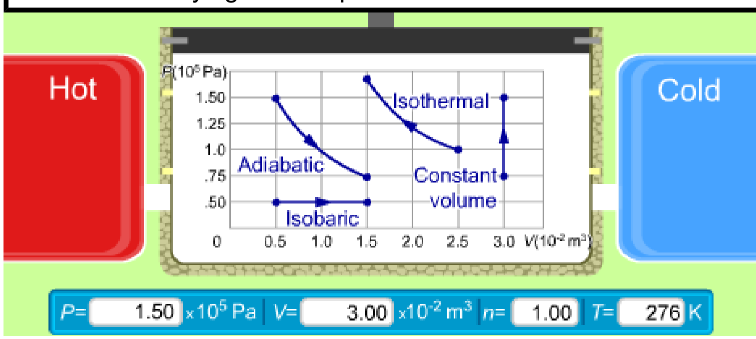
Work done depends on cycle
Cycle 2 does more work than Cycle 1

Above, you see two pressure-volume graphs. In both cases, the gas completes an entire cycle, starting and ending at the same pressure-volume point of 1.00×10^5 Pa and 2.00×10^{-2} cubic meters. Since in an engine cycle the gas returns to its initial state, the initial and final temperatures of the gas are the same as well.

However, as the two graphs above reflect, the work done by the gas differs. The **net** work done by the gas in each cycle is shown as a shaded area in the diagram. Each shaded area equals the positive work done by the gas when it expands (graph goes from left to right) plus the negative work that occurs when the gas contracts (graph goes from right to left). More work occurs during Cycle 2. We use these graphs to illustrate that the amount of work done by the engine depends on the nature of its cycle.

Why does more work occur during the cycle on the right? The gas expands with greater pressure in Cycle 2 than in Cycle 1. This occurs because more heat is added in Cycle 2. Since the gas in Cycle 2 supplies more force as it expands by the same amount as in Cycle 1, it does more work. By applying the first law of thermodynamics (or conservation of energy), you can confirm that more net heat energy is transferred into the container during Cycle 2, since the internal energy has not changed but the work done during that cycle is greater.

21.9 - Classifying thermal processes



concept 1

Gas cycles
Four types of gas processes:
· Constant volume: isochoric
· Constant pressure: isobaric
· Constant temperature: isothermal
· No heat transfer: adiabatic

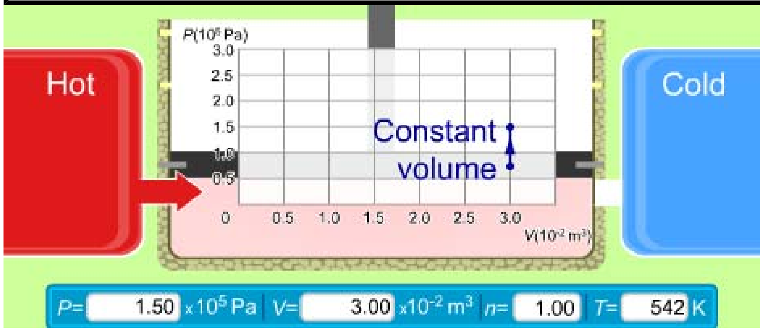
As a heat engine operates, the gas expands or contracts and its pressure changes. A gas can move from one pressure-volume state to another by different thermal processes. Four thermal processes are commonly studied and used in engines. These processes are characterized as follows.

1. Does the gas's volume remain constant? When it stays the same, the process is called a **constant-volume** process. This is also called an *isochoric process*.
2. Is the gas's pressure constant? If so, the process is called **isobaric**.
3. Does any heat flow in or out of the engine during the process? If there is **no** heat flow, the process is called **adiabatic**.
4. Does the temperature of the gas stay the same? If it does, the process is called **isothermal**.

In the following sections, we study each of these processes in more detail, based on their defining characteristics.

Above, we show each of these processes on a pressure-volume graph. Although we use an arrow to show each process occurring in a particular direction, the arrows could be reversed without changing the types of the processes. Any of these processes can occur in either direction.

21.10 - Constant-volume processes



concept 1

Constant-volume process
No movement, so no work

Constant-volume process: A process in which the volume of the gas remains the same.

In a heat engine like the one you see above, a constant-volume process results when the piston is locked in place. In a constant-volume process, the heat flow only changes the internal energy of the gas. You see this in Equation 1: The net heat transferred equals the change in the gas's internal energy.

This equation is not new. It is just a special case of the first law of thermodynamics. Since no work occurs, the net flow of heat equals the change in internal energy. When the net flow of heat is into the engine, the change in internal energy is a positive value. When the net flow is out of the engine, the change in internal energy is a negative value.

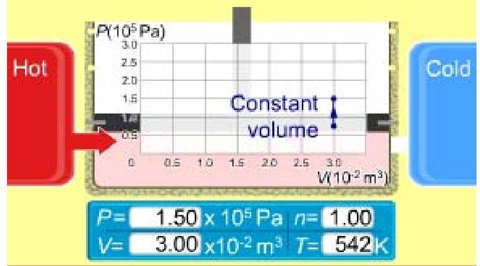
The change in internal energy changes the temperature of the gas. With the volume and the amount of gas held constant, the pressure changes proportionally to the temperature change, in accordance with the ideal gas law.

In an internal combustion engine, the ignition step of the cycle is a constant-volume process. The engine does no work during this process, but the heat transfer increases the internal temperature (and pressure) of the gasoline-air mixture that is the working substance of the engine. This energy is used during another process of the cycle to do work.

Example 1 shows a constant-volume process. This is the first process in an engine cycle we will use as an ongoing example to study thermal processes. In this process, heat is transferred to the gas at a constant volume. The engine is being prepared for a gas expansion process.

The processes being studied here might be found in a toy or model engine. The volume of the container, the amount of energy and the temperature ranges would be much larger in engines used in industry.

equation 1

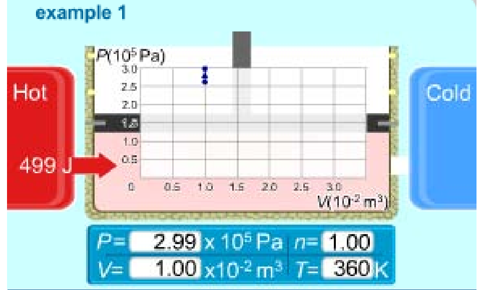


Constant-volume process
In a constant-volume process

$$Q = \Delta E_{\text{int}}$$

Q = net flow of heat
ΔE_{int} = change in internal energy

example 1



During this constant-volume process, 499 J of heat is transferred to the gas. What is the increase in the internal energy of the gas? How much work occurs?

$Q = \Delta E_{\text{int}}$

$\Delta E_{\text{int}} = Q = 499$ J

No work occurs

21.11 - Molar specific heat: constant volume

Molar specific heat, constant volume: A value that relates amount of heat to the change of temperature per mole of gas at a constant volume.

In a previous chapter, we stated an equation, $Q = kn\Delta T$, that relates the heat transferred to a given molar amount of a substance to its change in temperature. In this equation, k is a constant, the molar specific heat of the substance. For solids and liquids, as a good approximation, the constant k depends only on the substance. This is **not** true for gases. To measure specific heats for gases, either the volume is held constant or the pressure is held constant. Under either of these circumstances, a constant value for specific heat is obtained.

In this section, we examine the relationship of heat and temperature in a process during which the gas's volume is held constant. The first equation in Equation 1 relates the heat flow to the change in temperature. As you can see from the equation, the relationship is linear, as it is for a solid or liquid.

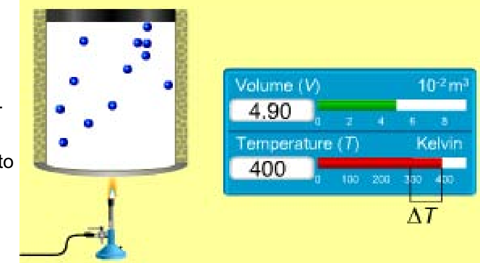
Let's analyze the physical reasons why the relationship should be linear. In a constant-volume process, no work is done, so the heat flow equals the gas's change in internal energy. We use this fact to write the second equation on the right, setting ΔE_{int} equal to the right side of the equation for Q.

We now apply two principles from the kinetic theory of gases. First, the temperature of the gas increases linearly with its internal energy, and since the change in internal energy equals the heat flow, it follows that temperature increases linearly with heat flow during a constant-volume process.

The second principle states that the internal energy E_{int} of an ideal monatomic gas is $(3/2)nRT$. Using that equation and the value of R, we can perform the steps below to calculate C_V for an ideal monatomic gas. Similar calculations can be done for diatomic and polyatomic gases. The table in Equation 3 shows molar specific heats of some gases at constant volume, as measured in laboratories. The values depend on the type of gas. The table shows how close the actual values are to the ideal values, especially for monatomic (argon) and diatomic (hydrogen, nitrogen and oxygen) molecules.

Step	Reason
1. $\Delta E_{int} = nC_V\Delta T$	molar specific heat, constant volume
2. $\frac{3}{2}nR\Delta T = nC_V\Delta T$	internal energy of monatomic gas
3. $C_V = \frac{3}{2}R = 12.5\ \text{J/mol·K}$	solve for C_V, use gas constant

equation 1



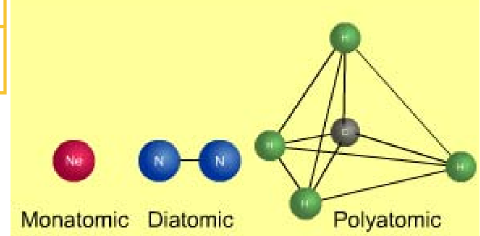
Molar specific heat, constant volume

$Q = nC_V\Delta T$

$\Delta E_{int} = nC_V\Delta T$

Q = heat, n = moles of gas
C_V = molar specific heat, constant volume
ΔT = change in temperature
ΔE_{int} = change in internal energy

equation 1



Molar specific heats, constant volume, for ideal gases

Monatomic: $(3/2)R$ = 12.5 J/mol·K
Diatomic: $(5/2)R$ = 20.8 J/mol·K
Polyatomic: $3R$ = 24.9 J/mol·K

equation 2

Gas	Molar specific heat, constant volume (J/mol · K)
Argon (Ar)	12.5
Hydrogen (H_2)	21.6
Nitrogen (N_2)	20.8
Oxygen (O_2)	21.1
Methane (CH_4)	27.5
Propane (C_3H_8)	66.2

at 1 bar, 300 K

Molar specific heats, constant volume

example 1

2.0 moles of nitrogen gas (N_2) increase 58 K in temperature at a constant volume. What is the

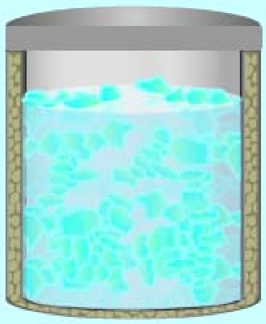
increase in internal energy of the gas?

$\Delta E_{int} = nC_V \Delta T$

$\Delta E_{int} = (2.0\text{ mol})(20.8\text{ J/mol·K})(58\text{ K})$

$\Delta E_{int} = 2400\text{ J}$

21.12 - Sample problem: melting ice



A well-insulated container holds 0.0230 moles of a monatomic gas at 32.0°C, as well as a mixture of crushed ice and water at 0°C. Some of the ice melts as the system reaches thermal equilibrium. How much heat transfers out of the gas? What mass of ice melts?

Assume the change in volume of the gas is negligible and treat this as a constant-volume process.

Variables

moles of gas	$n = 0.0230$ mol
initial gas temperature	$T_i = 32.0°$ C
final gas temperature	$T_f = 0°$ C
heat transferred	Q
internal energy	E_{int}
latent heat of fusion for water	$L_f = 3.34\times10^5$ J/kg

What is the strategy?

1. Use the energy equation for an ideal gas to calculate the change in internal energy. Since this is a constant-volume process, it equals the heat transferred from the gas.
2. Calculate the mass of ice melted from the latent heat of fusion for water.

Physics principles and equations

Heat transfers from the gas to the ice/water mixture. The heat transferred, Q, will be negative when we calculate it as being transferred out of the gas. We will use the corresponding positive value when we consider it being transferred into the ice/water mixture to reflect the direction of the transfer.

In a constant-volume process, heat transferred equals the change in internal energy.

$$Q = \Delta E_{int}$$

The internal energy for an ideal monatomic gas is

$$E_{int} = \frac{3}{2}nRT$$

The latent heat equation

$$Q = L_f m$$

Step-by-step solution

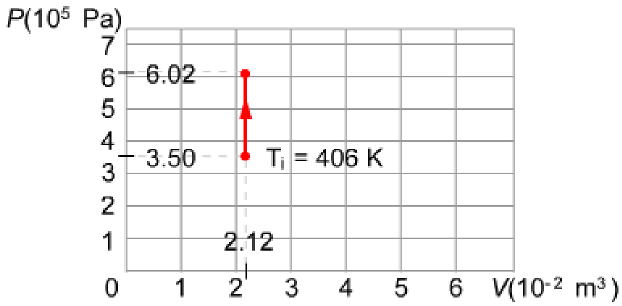
We start by computing the heat transferred out of the gas, which is the same as the change in internal energy of the gas.

Step	Reason
1. $Q = \Delta E_{int}$	constant-volume process
2. $E_{int} = \frac{3}{2}nRT$	internal energy of monatomic gas
3. $\Delta E_{int} = \frac{3}{2}nR(T_f - T_i)$	definition of change
4. $\Delta E_{int} = \frac{3}{2}(0.0230\text{ mol})(8.31\text{ J/mol}\cdot\text{K})(0°\text{C} - 32°\text{C})$	enter values
5. $Q = -9.17$ J	from steps 1 and 4

Now we can use the latent heat equation to calculate how much ice melted. The amount of (negative) heat that was transferred from the gas is the same as the amount of (positive) heat that is absorbed by the ice/water mixture.

Step	Reason
6. $Q = L_f m$	latent heat equation
7. $m = (9.17 \text{ J}) / (3.34\times10^5 \text{ J/kg})$	enter values
8. $m = 2.75\times10^{-5}$ kg	evaluate

21.13 - Interactive checkpoint: a constant-volume process



A quantity of ideal diatomic gas undergoes the thermal process shown. What is the final temperature of the gas? What amount of heat is transferred to the gas?

Answer:

$T_f =$ ______ K

$Q =$ ______ J

21.14 - Degrees of freedom

The molar specific heats at constant volume (C_V) for monatomic, diatomic and polyatomic gases are approximately $(3/2)R$, $(5/2)R$ and $3R$ respectively.

The internal energy of an ideal gas is solely a function of its temperature. Based on the molar specific heat constants above, a mole of ideal monatomic gas would have an internal energy of $(3/2)RT$, a diatomic gas would have $(5/2)RT$, and so on. Using the relationship between k and R, it follows from this that the atoms that make up a monatomic gas have an average energy of $(3/2)kT$, and the molecules of diatomic and polyatomic molecules have average energies of $(5/2)kT$ and $3kT$ respectively.

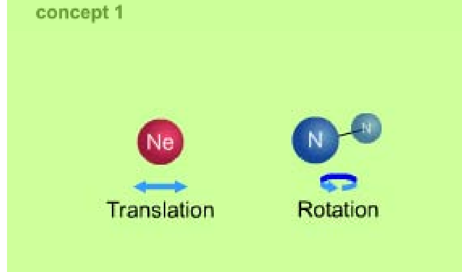
Why is this so? Why does the molecular form of the gas determine the relationship between its energy and temperature? James Maxwell's *equipartition of energy theorem* supplies an explanation. Maxwell asserted that atoms and molecules have a certain number of degrees of freedom, independent ways to store energy. For each degree of freedom, an atom or molecule on average has energy $(1/2)kT$ and one mole of gas has energy $(1/2)RT$.

The theorem is premised on the various ways in which an atom or molecule can store energy. Kinetic energy can be translational (linear) or rotational. Any atom or molecule can move in the x, y, or z direction, molecules can rotate around axes in some or all of these directions, and each direction represents a degree of freedom. Each of these three degrees of freedom accounts for $(1/2)kT$ of translational kinetic energy for an atom or molecule.

The atoms that make up a monatomic gas have only translational kinetic energy. They have no rotational kinetic energy. Why? Diatomic and polyatomic molecules have axes to rotate about, as shown in Concept 1, but an individual atom has no such axis. The individual atoms do not have rotational kinetic energy. This is a conclusion of quantum theory. The same conclusion can be drawn by considering the atoms as point masses. A point mass has a moment of inertia of zero, and can have no rotational kinetic energy. Since single atoms only have three degrees of freedom, monatomic gases should have a molar specific heat of $(3/2)R$, and this value is close to the actual measured values.

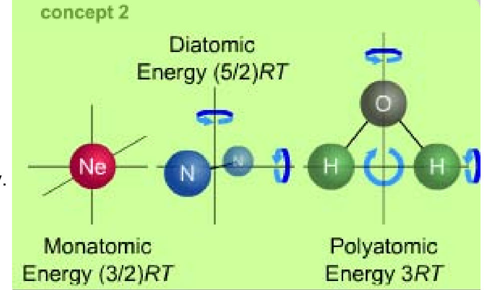
Diatomic and polyatomic molecules can rotate. A diatomic molecule rotates about the two axes perpendicular to the line connecting the two molecules, as shown in Concept 2. This means a diatomic molecule has two rotational degrees of freedom in addition to its three translational degrees for a total of five degrees of freedom. This accords with

concept 1



Equipartition of energy theorem
Molecular motion has degrees of freedom:
· Independent ways to store energy
Energy stored each way is the same:
· average ½kT per molecule
· ½RT per mole of gas

concept 2



Degrees of freedom

the molar specific heat of $(5/2)R$ for diatomic gases.

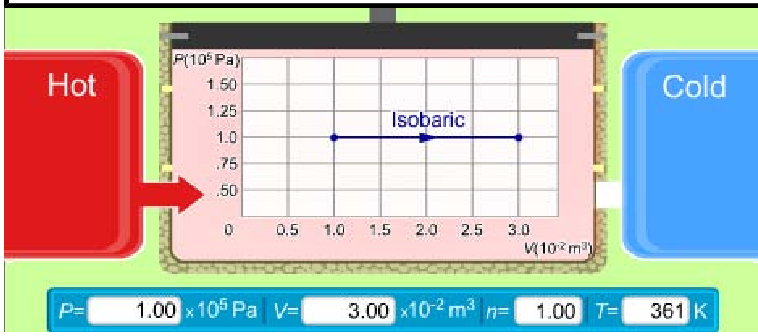
Polyatomic molecules have three degrees of rotational freedom. To understand these three degrees, imagine there is a central axis point in the polyatomic molecule in Concept 2. The molecule can rotate in all three dimensions about that point. This means a polyatomic molecule has a total of six degrees of freedom. This justifies a molar specific heat of $3R$. (One notable exception is carbon dioxide, CO_2 whose three atoms lie in a straight line; this gas has a molar specific heat closer to that of a diatomic gas.)

Monatomic: 3 degrees, $C_V = (3/2)R$
Diatomic: 5 degrees, $C_V = (5/2)R$
Polyatomic: 6 degrees, $C_V = 3R$

Maxwell's theorem has its limitations, however, and quantum theory supplies a more complete explanation of internal energies. For instance, the molar specific heat of a gas does change as it moves from extremely low to extremely high temperatures. This can be in part explained by considering other forms of energy, such as vibrational energy.

In general, quantum theory explains how the degrees of freedom of an atom or molecule change as temperature changes. The theory also explains why Maxwell's equipartition of energy theorem supplies a good approximation but not the exact values for the molar specific heat by providing a model for all the ways in which energy can be stored.

21.15 - Isobaric processes



concept 1

Isobaric process
Constant pressure

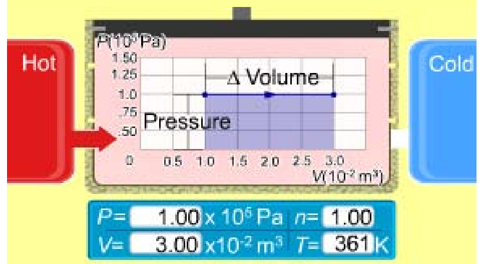
Isobaric process: Expansion or contraction during which pressure remains constant.

Above, you see an illustration of an isobaric process, one in which the gas pressure does not change. Heat transfers to the gas, the gas expands and its temperature increases proportionally, but the pressure does not change.

The work done by the gas is the area under the line on a pressure-volume graph of the process. Since the pressure is constant, the line is horizontal and the work equals the constant pressure at which the process occurs times the change in volume. We show this in Equation 1.

In Example 1, we calculate the work done by the gas during the second process in an engine cycle that began with a constant-volume process. The second process is an isobaric one in which the gas does a slight amount of positive work as it pushes the piston a short distance upward.

equation 1

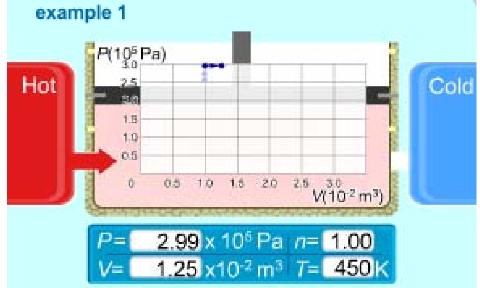


Isobaric process

$$W = P\Delta V$$

W = work done by gas
P = pressure
V = volume

example 1



The second process is isobaric, with pressure 2.99×10^5 Pa. The volume increases from 1.00×10^{-2} m^3 to 1.25×10^{-2} m^3. How much work is done by the gas?

$W = P\Delta V$
$\Delta V = 1.25\times10^{-2} - 1.00\times10^{-2}\ \text{m}^3$
$\Delta V = 0.25\times10^{-2}\ \text{m}^3$
$W = (2.99\times10^5\ \text{Pa})(0.25\times10^{-2}\ \text{m}^3)$
$W = 750\ \text{J}$

21.16 - Molar specific heat: constant pressure

Molar specific heat, constant pressure: A value that relates amount of heat to the change of temperature per mole of a gas at a constant pressure.

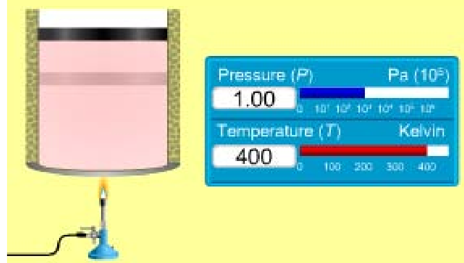
Molar specific heat relates heat transfer and temperature change for a given number of moles of a substance. Here, we discuss this value for a gas that is expanding at a constant pressure. The first equation in Equation 1 to the right shows the relationship between heat and temperature change when the pressure is constant.

The second equation states that the molar specific heat for a constant pressure process is greater than the molar specific heat for the same gas in a constant volume process. Why is that? A higher molar specific heat means that for the same transfer of heat into a gas, the temperature rise is not as great. In a constant pressure process, we add heat to a gas and allow the piston to move (which is necessary, otherwise its pressure would rise). The heat energy we put in only goes **partly** toward increasing the translational KE of the molecules, which is what we measure as a temperature rise. By the first law of thermodynamics, the other part of the heat energy goes to doing work. These two uses of the heat energy mean the molar specific heat at constant pressure is **greater** than the molar specific heat at constant volume.

We do not need to distinguish the two types of molar specific heat for liquids and solids. Since the amount of expansion of these phases of matter is relatively negligible for typical changes in temperatures, the work done by liquids and solids during expansion is negligible.

For an ideal gas, the molar specific heat coefficients at constant pressure and at constant volume are related by the second equation in Equation 1. The constant **pressure** molar specific heat equals the constant **volume** molar specific heat plus the gas constant, R. This equation provides accurate estimates for C_P for monatomic, diatomic and polyatomic gases at low density. In Equation 2, you see a table of molar specific heats for some gases at constant pressure.

equation 1



Molar specific heat, constant pressure

$Q = nC_P\Delta T$

$C_P = C_V + R$ (for ideal gas)

Q = heat, n = number of moles
ΔT = change in temperature
C_P = molar specific heat, constant pressure
C_V = molar specific heat, constant volume
R = gas constant

equation 2

Gas	Molar specific heat, constant pressure (J/mol · K)
Argon (Ar)	20.8
Hydrogen (H_2)	29.9
Methane (CH_4)	35.9
Nitrogen (N_2)	29.2
Oxygen (O_2)	29.4
Propane (C_3H_8)	75.1
	at 1 bar, 300 K

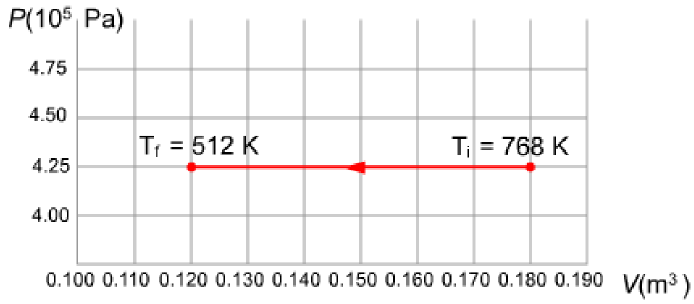
Molar specific heats, constant pressure

example 1

How much heat must be added to 2.0 moles of nitrogen gas (N_2) in an isobaric process to raise its temperature 67 K?

$Q = nC_P\Delta T$
$Q = (2.0\ \text{mol})(29.2\ \text{J/mol·K})(67\ \text{K})$
$Q = 3900\ \text{J}$

21.17 - Interactive checkpoint: a process with hydrogen

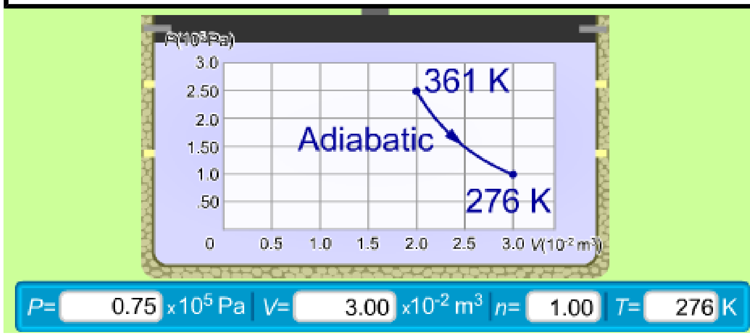


12.0 mol of hydrogen gas (H_2) undergoes the thermal process shown in the graph. What is its change in internal energy?

Answer:

$\Delta E_{\text{int}} =$ ________ J

21.18 - Adiabatic processes



concept 1

Adiabatic process
No exchange of heat

Adiabatic process: Expansion or contraction of a gas where there is no heat exchange between the system and its environment.

In an adiabatic process, a gas expands or contracts without the exchange of heat.

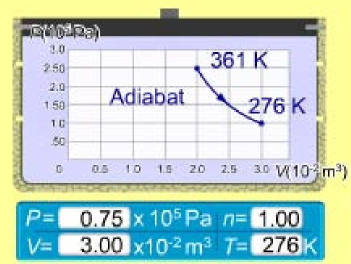
The illustration above shows an adiabatic system. The container is perfectly insulated, preventing any heat transfer. If the gas does positive work, raising the piston, that energy must come from the internal energy of the gas. No heat can flow in to supply the energy to do that work.

Insulating a system well is one manner in which to accomplish an adiabatic process. Extremely rapid processes can be adiabatic, as well, since there is too little time for significant heat transfer to occur. This is true for the expansion cycle in an internal combustion engine, and for the expansion and contraction that occurs with sound waves.

The first law of thermodynamics can again be used to analyze this process. Since no heat transfers into the gas, the work done by the gas on the piston equals the negative of its change in internal energy. You see this as the first equation in Equation 1. The path shown on the graph is called an *adiabat*. Positive work done by the gas decreases its internal energy, which accounts for the negative sign in the equation. A system that does positive work is a cooling system. Conversely, if the gas is compressed, the work done by the gas is negative and the change in its internal energy is positive.

For an example of this, blow a thin stream of air onto the back of your hand. The air expands adiabatically as it leaves your mouth, and it cools significantly compared to air that you exhale with your mouth wide open.

equation 1



Adiabatic process

$W = -\Delta E_{\text{int}}$

Ideal monatomic gas:

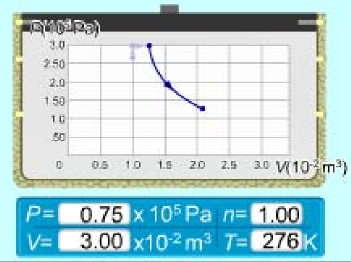
$W = -\frac{3}{2} nR(T_f - T_i)$

W = work done by gas
E_{int} = internal energy
n = moles, R = gas constant
T = temperature

The internal energy of an ideal monatomic gas is $(3/2)nRT$ (this equation was stated in the discussion of the kinetic theory of gases). Combining this equation with the first equation on the right allows us to derive the second equation shown, which relates work to the change in temperature. Again note the sign: A decrease in temperature means the gas does work.

In Example 1, we use the second equation to find the work done in the third process of an engine cycle. Heat was added in the previous isobaric process. Now the gas continues to expand, and do work, although no heat is transferred to the gas. The temperature of the gas returns to its initial value at the beginning of the four-process cycle, 320 K.

example 1



In the third phase of an engine cycle, an ideal monatomic gas expands adiabatically as its temperature changes from 450 K to 320 K. How much work does it do?

$W = -\frac{3}{2}nR(T_f - T_i)$

$T_f - T_i = 320\text{ K} - 450\text{ K} = -130\text{ K}$

$W = -\frac{3}{2}(1\text{mol})(8.31\text{J/mol}\cdot\text{K})(-130\text{ K})$

$W = 1620\text{ J}$

21.19 - Adiabatic processes, ideal gas

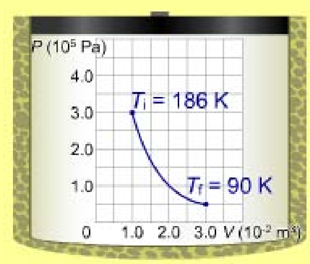
In the previous section, we related the work done by an ideal monatomic gas in an adiabatic process to the change in the gas's temperature. In this section, we discuss some equations that summarize the relationship of pressure, volume and temperature in an adiabatic process.

The equations in Equation 1 show the relationship between pressure and volume for an ideal gas in an adiabatic process. In these equations, the pressure P is multiplied by the volume V, raised to the power of γ (the Greek letter gamma). The exponent γ equals the ratio of the molar specific heat at constant pressure to the molar specific heat at constant volume. The product PV^γ is constant. Or, to state this as the second equation does: The pressure times volume raised to γ at a final state equals the pressure times volume raised to γ at the initial state. Because pressure is inversely related to a power of volume, the pressure-volume graph for an adiabatic process is a curve like the one you see in Equation 1.

For a constant amount of gas, the ideal gas law states that the pressure is proportional to the temperature divided by the volume. Using this fact, we can replace the pressure term with temperature, resulting in the pair of equations shown in Equation 2.

A table of molar specific heats for constant volume and constant pressure conditions, and the corresponding values of γ, is shown in Equation 3. You can use this table to solve problems like Example 1. Notice that for monatomic gases, γ is about 1.67. Why? For a monatomic gas, C_V is close to $(3/2)R$ and C_P is close to $(5/2)R$ (it equals $C_V + R$), so γ is approximately 5/3.

equation 1



Adiabatic process, ideal gas

$PV^\gamma = \text{constant}$

$P_f V_f^\gamma = P_i V_i^\gamma$

P = pressure, V = volume

$\gamma = C_P / C_V$

C_P = molar specific heat, constant pressure

C_V = molar specific heat, constant volume

equation 2

Adiabatic process, ideal gas

$TV^{\gamma-1} = \text{constant}$

$T_f V_f^{\gamma-1} = T_i V_i^{\gamma-1}$

T = temperature

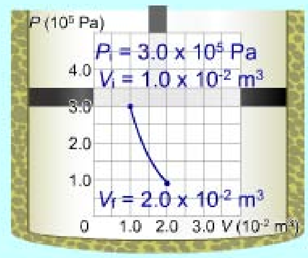
equation 3

Gas	C_P (J/mol · K)	C_V (J/mol · K)	$\gamma = C_P/C_V$
Argon (Ar)	20.8	12.5	1.66
Hydrogen (H_2)	29.9	21.6	1.38
Methane (CH_4)	35.9	27.5	1.31
Nitrogen (N_2)	29.2	20.8	1.40
Oxygen (O_2)	29.4	21.1	1.39
Propane (C_3H_8)	75.1	66.2	1.13

at 1 bar, 300 K

Molar specific heats and γ

example 1



$\gamma = 1.67$

Helium expands adiabatically as shown. What is its final pressure?

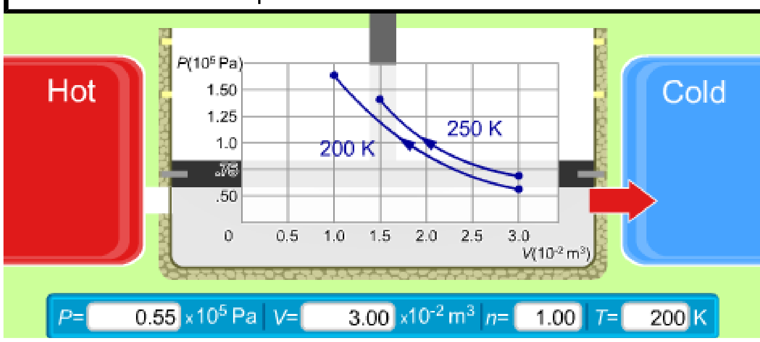
$$P_f V_f^{\gamma} = P_i V_i^{\gamma}$$

$$P_f = \frac{P_i V_i^{\gamma}}{V_f^{\gamma}}$$

$$P_f = \frac{(3.0 \times 10^5 \text{Pa})(1.0 \times 10^{-2} \text{m}^3)^{1.67}}{(2.0 \times 10^{-2} \text{m}^3)^{1.67}}$$

$$P_f = 9.4 \times 10^4 \text{ Pa}$$

21.20 - Isothermal processes



concept 1

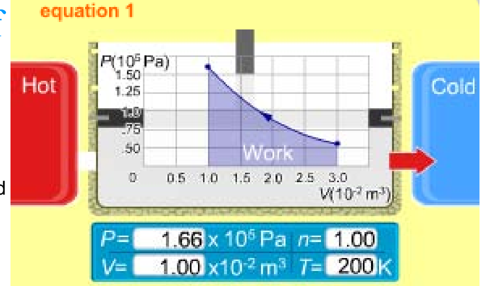
Isothermal process
Constant temperature
Heat equals work
Graph is reciprocal function, $P = k/V$

Isothermal process: Expansion or contraction of a gas with no change in its temperature.

An isothermal process is one in which the gas's temperature remains constant. Above, you see the pressure-volume graphs of two isothermal processes. The nature of these curves can be derived from the ideal gas law: $PV = nRT$. Since the temperature T (and the amount of gas n) remain constant, P is proportional to $1/V$. The curves shown represent inverse proportionalities; specifically, pressure-volume graphs at 200 K and 250 K with one mole of gas present.

During an isothermal process, the internal energy remains constant because the gas's temperature does not change. Applying the first law of thermodynamics, this means the net heat transferred to the gas equals the amount of work the gas does on the piston. If

equation 1



Isothermal process

the piston goes down, the gas contracts and it does negative work on the piston, expelling heat.

The amount of work done by the gas equals the area under the curve, taking into account the direction of the change in volume. The work is positive when the gas expands and negative when it contracts. Using calculus, we can derive the second equation in Equation 1 for the amount of work done by the gas. The temperature used in this equation must be measured in kelvins.

The equation states that the work is a constant value (nRT) times the natural logarithm of the gas's final volume divided by its initial volume. Logarithms are positive for numbers greater than one and negative for numbers less than one. These signs confirm that, as always, the gas does positive work when it expands and negative work when it contracts.

In Example 1, we compute the work done in the fourth and final process of the engine cycle we are analyzing. The process is isothermal. At the end of the previous adiabatic process, the temperature had returned to its initial value at the beginning of the cycle, 320 K. Now, the pressure and volume must also return to their initial values so the cycle can begin again. Since the volume decreases, the gas does negative work.

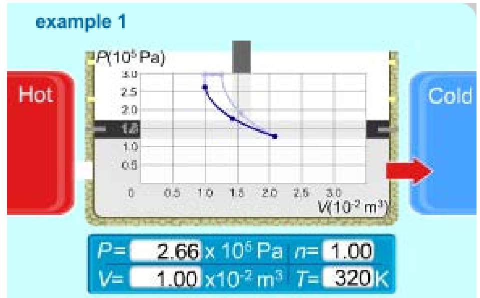
We can now review all four processes of the engine cycle we have been following in various sections of this chapter. You see the graph of the cycle in the illustration in Example 1.

1. First heat was transferred to the gas, but the piston was locked in place, so it was a constant-volume process. No work occurred during this cycle, but the gas's internal energy increased.
2. Next, we continued to transfer heat, and the gas was allowed to expand at a constant pressure, making this an isobaric process. The gas did positive work, raising the piston.
3. The third process was an adiabatic expansion: The gas did positive work on the piston but no heat was transferred. This means the gas's internal energy decreased. In our example, its temperature returned to its initial value.
4. Finally, in the **isothermal process** described in this section, the engine expels an amount of heat equal to the negative work done by the gas, so that its internal energy stays the same. In this last process, the gas's pressure and volume also return to their initial values and the engine cycle is complete.

$$W = Q$$

$$W = nRT \ln(V_f/V_i)$$

W = work, Q = net heat flow
n = moles of gas, R = gas constant
T = temperature (K), V = volume



In the last process of an engine cycle, the volume decreases from 2.09×10^{-2} m³ to 1.00×10^{-2} m³ isothermally at 320 K. How much work is done?

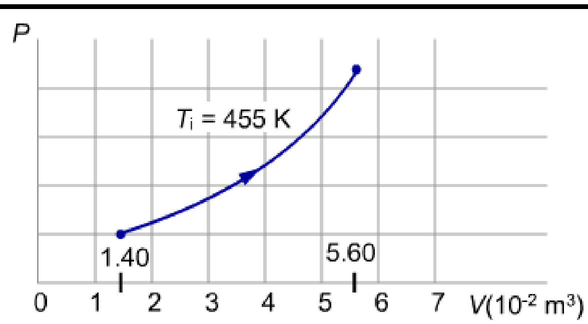
$$W = nRT \ln(V_f/V_i)$$

$$W = (1.00)(8.31)(320)\ln\frac{1.00\times10^{-2}}{2.09\times10^{-2}}$$

$$W = -1960 \text{ J}$$

How much work did the gas (or engine) do during this cycle? In previous example problems, we calculated the work done by the gas in the isobaric and adiabatic processes as 748 J and 1620 J, respectively. In the final isothermal process, −1960 J of work is done by the gas. The net work done by the engine is 408 J.

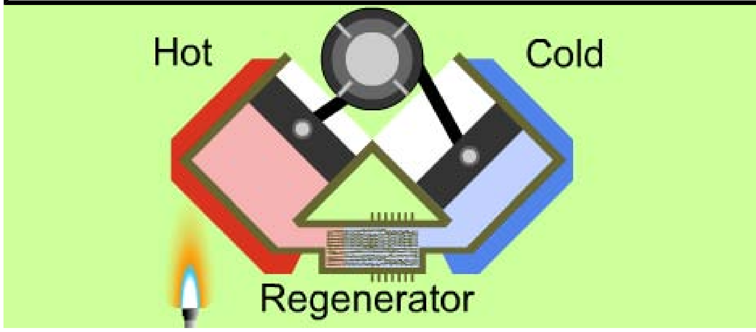
21.21 - Interactive checkpoint: a process with argon



3.40 mol of argon gas undergoes the thermal process shown in the graph. No heat exchange occurs during this process. How much work is done? Hint: First decide what kind of process this is.

Answer:

$W =$ ____ J



concept 1

Stirling engine
Performs work with fixed amount of gas
Uses regenerator for efficiency

Stirling engine: A heat engine designed to work extremely efficiently.

We have examined four thermal processes in this chapter, and put them together in a hypothetical engine cycle. Now, it is time to look at a real engine. Above, you see a diagram of a Stirling engine, an engine cycle conceived of almost 200 years ago by Robert Stirling, a Scottish minister. It is highly efficient, and in these days of environmental concern there is renewed interest in its design. Many websites are devoted to the Stirling engine.

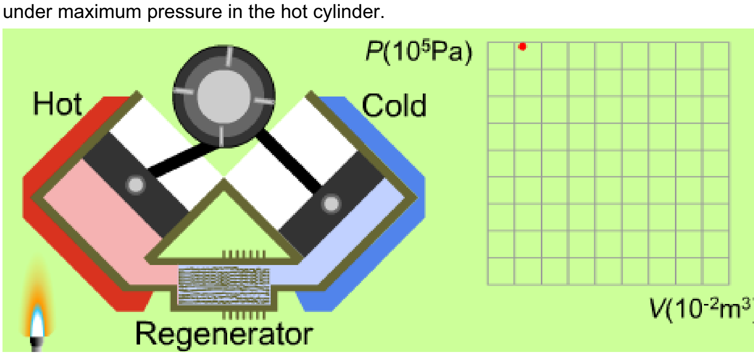
The Stirling engine works by expanding and compressing a fixed amount of gas in a closed container. We made the assumption when analyzing engine processes that the gas was uniform and had the same temperature and pressure everywhere. This assumption does not hold for the Stirling engine. Its operation depends on the gas having different temperatures in different parts of the engine. Nonetheless, two of the processes in a Stirling cycle are essentially isothermal, and two occur at constant volume and do no work. The heat source for a Stirling engine is external. A Stirling engine is an example of an external combustion engine, as is a steam engine.

The continued interest in Stirling engines stems from their efficiency. They can also draw their energy from a wide range of sources, from sunlight to the heat flow caused by a glass of ice water. However, they are not widely used because their power-to-weight ratio is not sufficient for most applications.

There are several variations in Stirling engine design. We will discuss a design called an "Alpha-type" engine, which uses two separate cylinders, one "hot" and one "cold", and two pistons. You see the cylinders and pistons in the illustration above. The cylinders act as the hot and cold reservoirs for the engine. A fixed amount of gas moves between the cylinders. This gas, distributed between the two cylinders, is the working substance of the engine. The pistons are connected and move in and out of the cylinders, expanding and compressing the gas, and moving it from one side to the other. The pistons do **not** alternate positions. Either piston can be up, near the rotating central mechanism, or down, independently. We will ignore the exact workings of the mechanism that links the pistons.

The efficiency of a Stirling engine results from the clever use of a *regenerator* that sits between the hot and cold cylinders in the engine. The regenerator absorbs heat when it is not desired, increasing its own temperature, and then gives up heat to the gas during another part of the engine cycle when it is desired. Typical regenerators are made of metal mesh through which the gas passes as it moves from cylinder to cylinder. When hot gas moves to the cold cylinder, heat flows from the gas to the regenerator. When cooled gas moves back, heat flows from the regenerator back into the gas.

The engine functions as follows. We start with the engine at the beginning of a cycle, with both pistons down, near the regenerator, and the gas under maximum pressure in the hot cylinder.

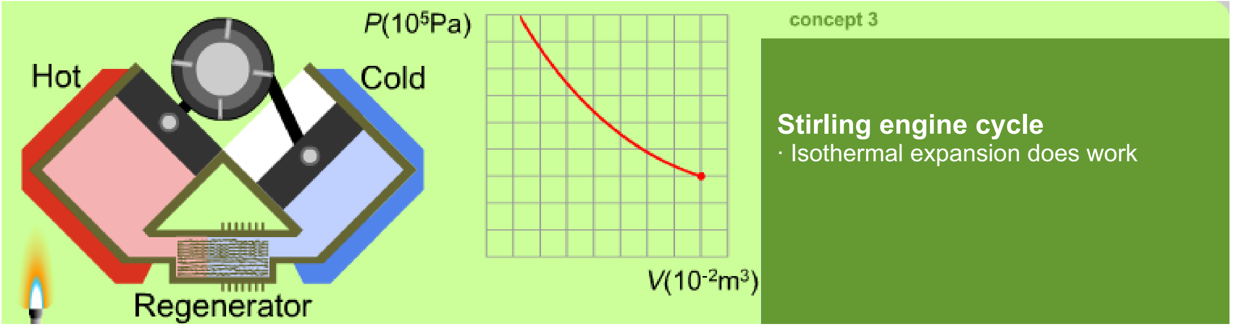


concept 2

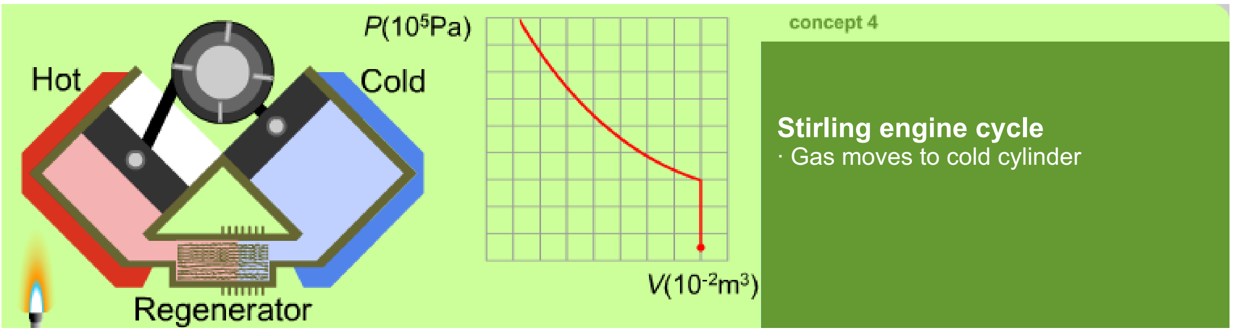
Stirling engine cycle

In the steps below, the illustrations show the positions of the pistons at the **end** of each step. We also draw the pressure-volume graph at each step. The volume is the total volume of the gas in both cylinders.

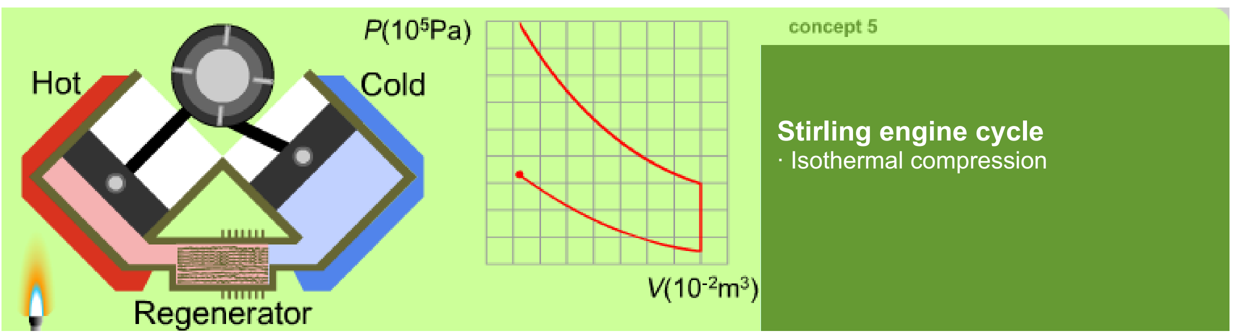
1. As the gas gains heat from the hot cylinder, it expands, forcing the hot piston out while the cold piston moves only slightly. This is an **isothermal expansion** in the hot cylinder, meaning the temperature of the gas there does not change. This is the step of the cycle where the gas does positive work, pushing the piston in the hot cylinder upward.



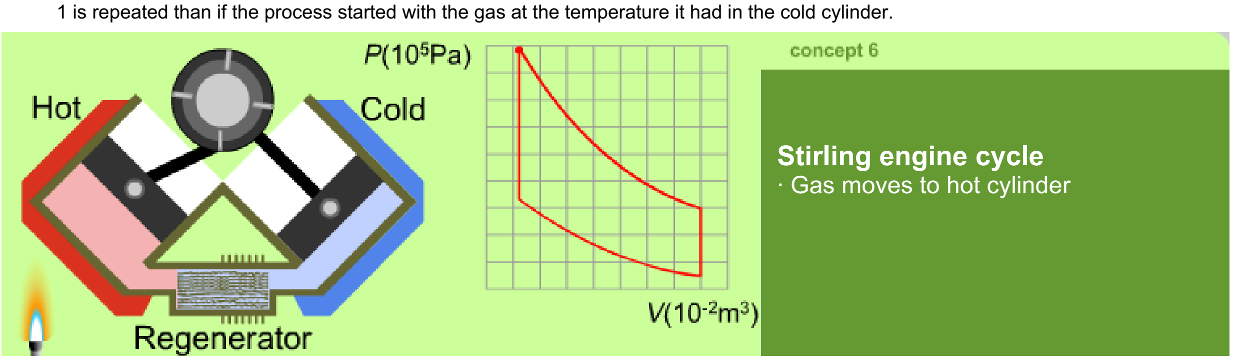
2. The mechanism causes the hot piston to move down and the cold piston to move up at the same time, forcing the gas through the regenerator into the cold cylinder. The pistons move in synchrony so that the volume of the gas is constant. No work occurs in this step. The regenerator absorbs some of the heat, reducing the temperature of the gas moving to the cold cylinder. This means the cold cylinder will have less heat to dissipate.



3. The cold piston now moves down, compressing the gas, while the hot piston moves only a little. Heat transfers from the gas to the cold cylinder. This is an **isothermal compression** in the cold cylinder.



4. The pistons both move to shift the gas back into the hot cylinder. Again, the gas volume is constant during this step and no work occurs. The gas gains some heat back from the regenerator as it passes by, and increases in temperature. The cycle can now begin again. Since the regenerator has increased the gas's temperature, less heat has to be transferred from the walls of the hot cylinder when step 1 is repeated than if the process started with the gas at the temperature it had in the cold cylinder.



There are other types of Stirling engines. Some use a "displacer" instead of a second piston. For example, you can build or buy small Stirling engines that run on the heat of a coffee cup, or even the warmth of your hand. In these engines, a metal mesh displacer that doubles as the regenerator shuffles the gas between the hot and cold sides of a broad shallow cylinder, but does not behave like a piston.

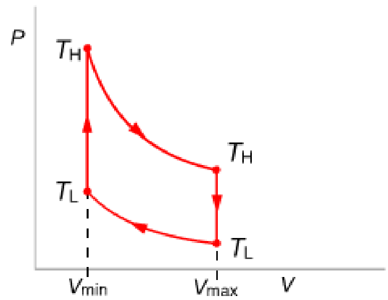
21.23 - Sample problem: Stirling engine and work

As it sits on the top of a cup of hot coffee, the temperature of 1.40×10^{-3} mol of ideal gas in the little Stirling engine ranges between 110°C and 120°C, and the volume between 7.93×10^{-5} m^3 and 8.01×10^{-5} m^3. How much work does the engine do in one cycle?

A Stirling engine has two constant-volume steps, during which no work occurs, and two isothermal steps. The isothermal expansion and compression steps are the only ones that contribute to the work. The net work done by the engine equals the sum of the work the gas does during these two steps.

Diagram

A pressure-volume graph of a Stirling engine cycle is helpful. We use T_H to indicate the highest temperature, T_L the lowest temperature, and V_{min} and V_{max} to indicate the minimum and maximum volumes.



Variables

high gas temperature	$T_H = 120°C = 393$ K
low gas temperature	$T_L = 110°C = 383$ K
minimum volume	$V_{min} = 7.93\times10^{-5}$ m^3
maximum volume	$V_{max} = 8.01\times10^{-5}$ m^3
number of moles of gas	$n = 1.40\times10^{-3}$ mol
work done by gas during expansion	W_{exp}
work done by gas during compression	W_{cmp}
net work	W

What is the strategy?

1. Use the equation for the work done by an isothermal process to write expressions for the work done during the isothermal expansion and compression steps. Add these expressions and simplify.
2. Enter the values and calculate the net work.

Physics principles and equations

The work done by a gas in an isothermal process is given by

$$W = nRT \ln(V_f/V_i)$$

Mathematical principles

The logarithm of a quotient is

$$\ln(a/b) = \ln(a) - \ln(b)$$

Step-by-step solution

We start by deriving an equation for the work done during one cycle of a Stirling engine. Only the isothermal phases contribute to the work.

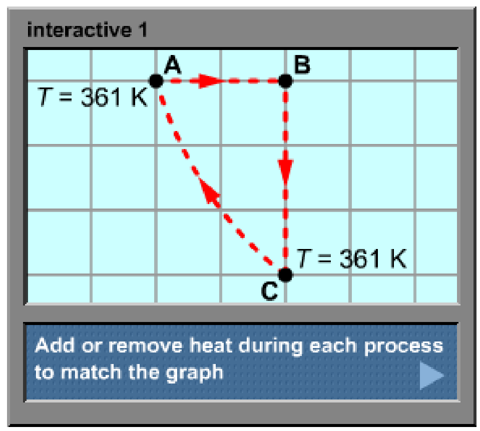
Step	Reason
1. $W = W_{exp} + W_{cmp}$	net work is sum
2. $W_{exp} = nRT_H \ln(V_{max}/V_{min})$	work done during expansion
3. $W_{exp} = nRT_H[\ln(V_{max}) - \ln(V_{min})]$	logarithm rule
4. $W_{cmp} = nRT_L[\ln(V_{min}) - \ln(V_{max})]$	work done during compression
5. $W = nR[(T_H - T_L)\ln V_{max} - (T_H - T_L)\ln V_{min}]$	substitute equations 3 and 4 into equation 1
6. $W = nR(T_H - T_L)\ln(V_{max}/V_{min})$	logarithm rule

Now we can enter the values given in the problem and calculate the amount of work done.

Step	Reason
7. $W = (1.40 \times 10^{-3})(8.31)(393 - 383) \ln \dfrac{8.01 \times 10^{-5}}{7.93 \times 10^{-5}}$	enter values
8. $W = 1.17 \times 10^{-3}$ J	evaluate

21.24 - Interactive summary problem: an engine cycle

A pressure-volume diagram for an engine cycle is shown in the interactive simulation. The cycle has three processes, and the points where one process ends and another starts are labeled A, B and C. The temperature at two points in the cycle is shown, and you can determine the pressure and volume at each point from the graph.



Your goal is to add or remove the proper amount of heat during each process to drive the system along the path shown. To do this, first determine the type of process (constant-volume, isobaric, adiabatic, or isothermal).

The engine contains **one mole** of an **ideal monatomic** gas; use this fact to determine the temperature not given to you for point B. You can then calculate the change in internal energy of the gas during each process and apply the appropriate equation for each process to determine the amount of work done. The first law of thermodynamics will then let you calculate the heat required for each process.

Enter the heat values for each process (use a negative value to indicate a net flow of heat out of the gas), to the nearest 10 joules, and press GO. If you have entered the correct values, the graph of the resulting engine cycle will match the given one. If not, press RESET to try again.

If you have trouble matching the target cycle, start by determining the type of each process. Then you should consider how, for each process, you can determine the work done, or the heat flow, or the initial and final temperatures. Do not forget the first law of thermodynamics, nor the fact that this is an ideal monatomic gas.

21.25 - Gotchas

In which process does all heat transferred in do work? Isothermal, since the temperature remains constant in this process.

In which process does all the energy for the work come from (or add to) the internal energy of the gas? An adiabatic process, since no heat is transferred in this process.

Does the gas do work in a constant-volume process? No, it does not, since the gas's volume does not change.

If you see a pressure-volume graph, you can determine how much work the gas does during a process. Yes, but be careful. The direction of the process matters. If the volume increases, the gas does positive work. If the volume decreases, the gas does negative work.

A gas is hotter after an engine cycle is complete. No, the definition of an engine cycle is that the engine returns to its initial state.

Adding heat to an engine immediately causes it to do work. Although that is likely the purpose of the heat transfer, the heat can also increase the internal energy of the engine. It may be later in the cycle that the engine uses the energy to do work.

Doing work on an engine can cause the temperature of its gas to increase. This is true. It can also cause heat to flow out of the engine as well.

The first law of thermodynamics is the same as the principle of conservation of energy. Essentially, this is true. The first law explicitly factors in work, a way to add or subtract energy from a system. However, both physics principles are statements of conservation. Energy cannot magically appear or disappear; it can always be accounted for.

James Joule demonstrated a fundamental relationship between heat, temperature and work. Yes. He showed how both work and heat could increase the temperature (and energy) of a substance.

21.26 - Summary

The first law of thermodynamics relates heat, work, and internal energy, and is a re-statement of the law of conservation of energy. It states that the net heat transferred to a system equals its change in internal energy plus the work done by the system.

A heat engine uses thermal energy to do useful work. An engine contains a working substance, usually a gas, that goes through a series of thermal processes. During a thermal process, heat can be transferred to the gas from a hot reservoir, or can flow out to a cold reservoir, or it could be that no heat exchange takes place. In an analogous way, during the process, energy transfers can also take place via work: work can be done on the system, work can be done by the system on its surroundings, or no work might be done by the system. An engine cycle is a series of processes that returns the gas to its original temperature, pressure, and volume. During a cycle, because the gas has returned to its original state and the change in its internal energy is zero, the net work done equals the difference between the heat transferred into the engine and the heat transferred out.

When the amount of gas is constant, the ideal gas law states that the product of pressure and volume is proportional to the temperature of the gas.

A pressure-volume graph is a valuable tool for understanding engine processes. The amount of work done by a gas during a process is the area under its pressure-volume graph. The work is positive if the volume increases, and negative if the volume decreases.

Four thermal processes occur frequently in engines.

Constant-volume process: The volume of the gas does not change. A constant-volume process does no work, so the net heat transfer equals the change in the internal energy of the gas.

Isobaric process: The pressure does not change. The work done during an isobaric process can be calculated as the constant pressure times the change in volume.

Adiabatic process: No heat transfers to or from the gas. The work done during an adiabatic process is the negative of the change in the gas's internal energy.

Isothermal process: The temperature does not change. All the heat transferred in an isothermal process goes into work.

Using the molar specific heat of a gas, the change in its temperature for a given amount of heat transfer can be calculated. The molar specific heat is specified under conditions of constant volume or constant pressure. For an ideal gas, the molar specific heat at constant pressure equals the molar specific heat at constant volume plus the gas constant R.

Equations

First law of thermodynamics

$Q = \Delta E_{\text{int}} + W$

Engine cycle

$W = Q_h - Q_c$

$W = \int_{V_i}^{V_f} P\,dV$

Constant-volume process

$Q = \Delta E_{\text{int}}$

Molar specific heat, constant volume

$Q = nC_V \Delta T$

$\Delta E_{\text{int}} = nC_V \Delta T$

Monatomic: $(3/2)R = 12.5$ J/mol·K

Diatomic: $(5/2)R = 20.8$ J/mol·K

Polyatomic: $3R = 24.9$ J/mol·K

Isobaric process

$W = P\Delta V$

Molar specific heat, constant pressure

$Q = nC_P \Delta T$

$C_P = C_V + R$ (for ideal gas)

Adiabatic process

$W = -\Delta E_{\text{int}}$

$W = -\frac{3}{2} nR(T_f - T_i)$ (monatomic gas)

Adiabatic process, ideal gas

$PV^{\gamma} = \text{constant}$

$P_f V_f^{\gamma} = P_i V_i^{\gamma}$

$TV^{\gamma-1} = \text{constant}$

$T_f V_f^{\gamma-1} = T_i V_i^{\gamma-1}$

Isothermal process

$W = Q$

$W = nRT \ln(V_f/V_i)$

Chapter 21 Problems

Conceptual Problems

C.1 A gas is enclosed in a container whose walls are made of a material that is flexible and which acts as a perfect insulator for all practical purposes. A thermal process takes place on the gas. What would be the best model for the type of process?

i. Constant-volume
ii. Isobaric
iii. Adiabatic
iv. Isothermal

C.2 A gas is sealed in a container with rigid walls that allow heat to flow through them. A thermal process takes place on the gas. What would be the best model for the type of process?

i. Constant-volume
ii. Isobaric
iii. Adiabatic
iv. Isothermal

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) What is the relationship between the amount of heat transferred to the gas and the change in the internal energy of the gas during the first process when the volume is fixed? (b) How do the values for the work and change in internal energy relate during the second process when no heat flows in or out of the gas? (c) In general, how does the amount of heat transferred to the gas relate to the change in internal energy and the work it does?

(a) The heat is i. greater than / ii. equal to / iii. less than the change in internal energy.

(b) The work done is i. equal to / ii. the negative of the change in internal energy.

(c) The net heat transferred is i. greater than / ii. equal to / iii. less than the sum of the change in internal energy and the work done.

Section 1 - First law of thermodynamics

1.1 In the gym, you do a workout routine that causes you to lose 3.4×10^5 J of heat while performing 1.7×10^5 J of work. With yourself as the system, state the following quantities for the first law of thermodynamics, with the appropriate signs: (a) Q, (b) ΔE_{int}, (c) *W*.

(a) __________ J
(b) __________ J
(c) __________ J

1.2 A system performs 356 J of work on its surroundings while losing 133 J of internal energy. Determine the heat transferred to the system during this process.

__________ J

1.3 A sealed container of gas with a piston is well-insulated, so that no heat is transferred to it or from it. The gas performs 1700 J of work on its surroundings. What is its change in internal energy?

__________ J

1.4 An ideal monatomic gas is held in a well-insulated container with a piston, so that no heat is transferred to it or from it. There are 1.75 moles of gas in the container. During a process, the temperature of the gas increases from 297 K to 353 K. (a) What is the work associated with the process? Make sure to include the correct sign. (Hint: there is an equation in the chapter on the Kinetic Theory of Gases for the internal energy of an ideal monatomic gas that will help you solve this problem.) (b) Is work done on the gas, or does the gas do work on its suroundings?

(a) __________ J
(b) i. Work is done on the gas
ii. Work is done by the gas

Section 3 - Heat engines

3.1 During a cycle, a heat engine receives 2.78×10^4 J of heat from the hot reservoir and gives up 1.83×10^4 J of heat to the cold reservoir. (a) What is the net work performed during one cycle? Make sure to include the correct sign. (b) Is work done on the system, or does the system do work on its suroundings?

(a) __________ J
(b) i. Work is done on the system
ii. Work is done by the system

3.2 A heat engine does 2.4×10^5 J of work during each cycle. If the engine loses 1.8×10^5 J of heat to the cold reservoir in each cycle, how much heat is transferred from the hot reservoir during the cycle?

__________ J

3.3 A mountain climber gains 897 m in elevation on a particular hike. Consider the climber's body as a heat engine, with heat for work provided by burning food (the "hot reservoir") and excess heat radiated into the mountain air (the "cold reservoir"). (a) If her mass is 62.3 kg, how much work does she do changing her gravitational potential energy? (b) If she radiates 1.67×10^5 J of heat during the climb, and her internal energy doesn't change, how much heat does the food supply?

(a) __________ J
(b) __________ J

3.4 The horsepower rating commonly used for engines indicates the rate at which they can do work. A one-horsepower engine can do 746 J of work each second. A lawn mower has a 6.50 horsepower engine that runs at 3800 rpm. Each revolution of the engine represents one engine cycle. (a) What amount of work does the engine do on each cycle? (b) If the hot reservoir of the engine supplies 142 J of heat each cycle, hot much heat must the engine expel to the cold reservoir?

(a) __________ J
(b) __________ J

Section 7 - Work and pressure-volume graphs

7.1 During a particular process, the pressure of the gas in a heat engine increases from 1.15×10^5 Pa to 2.37×10^5 Pa, while the volume increases from 2.33×10^{-2} m^3 to 2.89×10^{-2} m^3. The pressure-volume graph is a straight line. (a) What is the work associated with the process? Make sure to include the correct sign. (b) Is work done on the system, or does the system do work on its suroundings?

(a) __________ J
(b) i. Work is done on the system
ii. Work is done by the system

7.2 During a particular process, the pressure of the gas in a heat engine increases from 1.62×10^5 Pa to 2.64×10^5 Pa, while the volume decreases from 3.34×10^{-2} m^3 to 2.74×10^{-2} m^3. The pressure-volume graph is a straight line. (a) What is the work associated with the process? Make sure to include the correct sign. (b) Is work done on the system, or does the system do work on its suroundings?

(a) __________ J
(b) i. Work is done on the system
ii. Work is done by the system

Section 10 - Constant-volume processes

10.1 A steel tank contains a large quantity of compressed air. When the tank is submerged for an hour in a cold-water tank, the air loses 245 J of heat to the water. What is the change in the internal energy of the air?

__________ J

10.2 A rigid aluminum cylinder contains some nitrogen gas. The cylinder has mass 0.935 kg, and it is in thermal equilibrium with the gas. After 2.41×10^4 J of heat are transferred to the system consisting of the cylinder and the gas, their temperature increases 27.9 K, and they are again in thermal equilibrium. What is the change in internal energy of the gas? (The specific heat of aluminum is 897 J/kg·K.)

__________ J

Section 11 - Molar specific heat: constant volume

11.1 A cylinder contains 3.5 mol of hydrogen gas at a constant volume. If the hydrogen gas absorbs 1500 J of heat, what is its change in temperature?

__________ K

11.2 How much heat must be added to 1.70 mol of nitrogen gas at 29.0°C to raise its temperature to 68.0°C, at the same volume?

__________ J

11.3 A rigid metal cylinder contains argon gas. The cylinder and the gas are in thermal equilibrium at 22°C, and are heated to 62°C, again in thermal equilibrium. In the process, 9700 J of heat are transferred to the cylinder and the gas, of which 7400 J are consumed raising the temperature of the cylinder. How many moles of gas are there in the cylinder?

__________ mol

11.4 A rigid iron cylinder contains propane gas. The cylinder and the gas are in thermal equilibrium. In a process, 8700 J of heat are transferred to the cylinder and the gas, so that at the end they are again in thermal equilibrium, with a temperature increase of 24 K. If the mass of the cylinder is 0.51 kg, how many moles of propane are there in the cylinder?

__________ mol

Section 15 - Isobaric processes

15.1 A gas is compressed from 4.5×10^{-2} m^3 to 2.6×10^{-2} m^3 at a constant pressure of 1.7×10^5 Pa. (a) What is the work associated with the process? Make sure to include the correct sign. (b) Is work done on the system, or does the system do work on its surroundings?

(a) __________ J

(b) i. Work is done on the system
ii. Work is done by the system

15.2 A gas performs 5600 J of work while it expands at a constant pressure of 2.1×10^5 Pa. If its initial volume is 0.037 m^3, what is its final volume?

__________ m^3

15.3 In a heat engine, 1800 J of heat are added to the gas at constant pressure, while its internal energy increases 3300 J and its volume decreases from 3.2×10^{-2} m^3 to 1.9×10^{-2} m^3. What is the constant gas pressure?

__________ Pa

15.4 In a heat engine, the working substance is an ideal monatomic gas. During a thermal process, the volume decreases from 2.80×10^{-2} m^3 to 1.30×10^{-2} m^3, while the pressure stays constant at 2.80×10^5 Pa. (a) What is the heat associated with the process? Make sure to include the correct sign. (An equation for the internal energy of an ideal monatomic gas will prove useful in solving this problem.) (b) Overall, is thermal energy added to the system, or does thermal energy leave the system?

(a) __________ J

(b) i. Thermal energy enters the system
ii. Thermal energy leaves the system

Section 16 - Molar specific heat: constant pressure

16.1 If 3.6 mol of methane is heated so that the temperature of the gas increases 45 K while its pressure remains constant, how much heat is added to the gas?

__________ J

16.2 6800 J of heat are added to 4.7 mol of methane at a constant pressure. The initial temperature of the methane is 38°C. What is its final temperature?

__________ ° C

16.3 During an isobaric process, the volume and the temperature of a large quantity of argon gas are related by the equation $V = T/120$, where V is measured in m^3 and T in K. If the volume of the gas decreases from 3.80 m^3 to 2.70 m^3, and the pressure is a constant 2.30×10^5 Pa, how many moles of argon are there?

__________ mol

16.4 Heat is added to the working substance of a heat engine, which in this case is nitrogen gas. You can assume the nitrogen is acting as an ideal gas. During an isobaric process, the gas does 3.80×10^3 J of work, while its volume expands from 2.70×10^{-2} m^3 to 4.20×10^{-2} m^3. There are 7.12 mol of gas. (a) What is the heat associated with the process? Make sure to include the correct sign. (b) Overall, is thermal energy added to the system, or does thermal energy leave the system?

(a) __________ J

(b) i. Thermal energy enters the system
ii. Thermal energy leaves the system

Section 18 - Adiabatic processes

18.1 An ideal monatomic gas undergoes an adabatic expansion, and its temperature changes from 194°C to 110°C. If the gas does 3580 J of work, how many moles of gas are there?

__________ mol

18.2 During an adiabatic process, the volume of an ideal monatomic gas changes from 0.340 m^3 to 0.540 m^3, while the pressure goes from 1.70×10^5 Pa to 7.86×10^4 Pa. The amount of gas is 2.44 mol. (a) What is the work associated with the process? Make sure to include the correct sign. (b) Is work done on the system, or does the system do work on its surroundings?

(a) __________ J

(b) i. Work is done on the system

ii. Work is done by the system

Section 19 - Adiabatic processes, ideal gas

19.1 A quantity of argon gas expands adiabatically as an ideal gas, from 4.57×10^{-2} m^3 to 5.92×10^{-2} m^3. If its initial temperature is 415 K, what is its final temperature?

__________ K

19.2 A quantity of methane, acting as an ideal gas, increases in temperature from 35.7°C to 58.5°C during an adiabatic process. Its final volume is 0.0358 m^3. What was its initial volume?

__________ m^3

19.3 During an adiabatic process, a quantity of nitrogen acting as an ideal gas decreases in volume from 0.0276 m^3 to 0.0142 m^3. Its initial pressure is 1.64×10^5 Pa. What is its final pressure?

__________ Pa

19.4 Propane, acting as an ideal gas, changes pressure from 2.23×10^5 Pa to 3.41×10^5 Pa. If the final volume of the gas is 0.0228 m^3, what was its initial volume?

__________ m^3

19.5 For any monatomic gas, γ is very close to 5/3. If a monatomic ideal gas is compressed adiabatically, so that its volume becomes half what it was initially, by what factor does its pressure increase? Express your answer to three significant digits.

____

19.6 The work done by 1.70 mol of an ideal monatomic gas during an adiabatic expansion is 9.30×10^2 J. The initial temperature of the gas is 420 K and its initial volume is 0.0120 m^3. Assume γ = 5/3 and calculate (a) the final temperature and (b) the final volume of the gas.

(a) __________ K

(b) __________ m^3

Section 20 - Isothermal processes

20.1 In an isothermal process, 6.90×10^3 J of heat is added to 1.45 mol of a gas. If the (constant) temperature of the ideal gas is 295 K, and the initial volume of the gas is 0.189 m^3, what is its final volume?

__________ m^3

20.2 As a quantity of ideal gas expands isothermally from 0.0180 m^3 to 0.0320 m^3, it does 5.80×10^3 J of work. If the amount of gas is 3.87 mol, what is the temperature of the gas?

__________ K

20.3 During an isothermal process at 324 K, 1.12 mol of an ideal gas decreases in pressure from 3.45×10^5 Pa to 2.74×10^5 Pa. (a) What is the work associated with the process? Make sure to include the correct sign. (b) Is work done on the system, or does the system do work on its suroundings?

(a) __________ J

(b) i. Work is done on the system

ii. Work is done by the system

20.4 Helium has an atomic mass of 4.00 u. Acting as an ideal gas, 3.20×10^{-3} kg of helium expand isothermally at 298 K, doing 8740 J of work in the process. If the initial volume of the gas is 0.0245 m^3, what is the final volume?

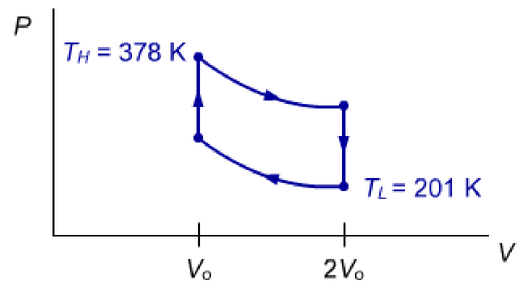
__________ m^3

Section 23 - Sample problem: Stirling engine and work

23.1 If a Stirling engine operates using 0.230 mol of an ideal gas, ranges in temperature between 394 K and 478 K, and ranges in volume between 0.0245 m^3 and 0.0352 m^3, how much work does it do during one cycle?

__________ J

23.2 Assume that an engine has been built, with an operating cycle somewhat similar to the Stirling cycle except that the two isothermal processes are replaced by adiabatic processes. (The two other steps take place at constant volume as in a Stirling cycle.) If such an engine operates using 0.162 mol of argon acting as an ideal gas, with a high temperature of 378 K and a low of 201 K, and the volume doubles during the adiabatic expansion, how much work does the engine do during one cycle?



__________ J

Section 24 - Interactive summary problem: an engine cycle

24.1 Use the information given in the interactive problem in this section to calculate the heat for (a) process A-B, (b) process B-C, and (c) process C-A, in order to match the path shown. Test your answer using the simulation.

(a) __________ J
(b) __________ J
(c) __________ J

Additional Problems

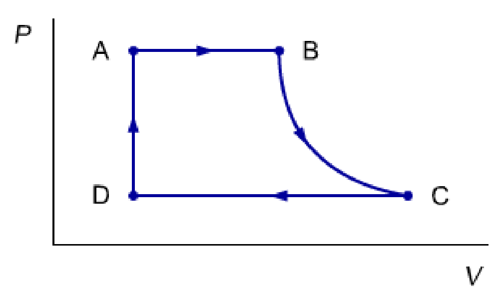
A.1 An ideal gas expands at a constant atmospheric pressure of 1.01×10^5 Pa, so that it doubles in volume. The gas is initially at 23.5°C, and the amount of gas is 2.13 mol. (a) What is the final temperature of the gas in kelvins? (b) How much work does the gas do?

(a) __________ K
(b) __________ J

A.2 At atmospheric pressure (1.01×10^5 Pa), 0.203 kg of liquid water at 100°C is boiled to produce 0.342 m^3 of steam, also at atmospheric pressure and 100°C. The density of liquid water at 100°C is 958.4 kg/m^3. (a) How much work is done by the water as it expands? (b) How much heat is added to the water during the process? (c) What is the change in its internal energy?

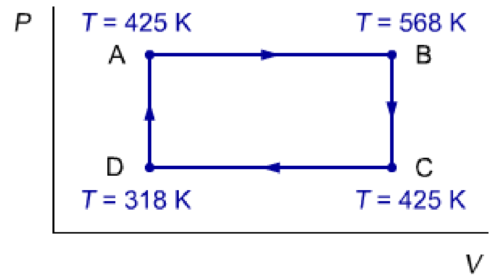
(a) __________ J
(b) __________ J
(c) __________ J

A.3 The accompanying pressure-volume graph shows an engine cycle with one isothermal process, B-C. The working substance is an ideal gas. At point B, the pressure is 3.20×10^5 Pa and the volume is 0.0480 m^3. At point C, the pressure is 2.10×10^5 Pa and the volume is 0.0730 m^3. At point A the volume is 0.0230 m^3. What is the net work done by the gas during one cycle?



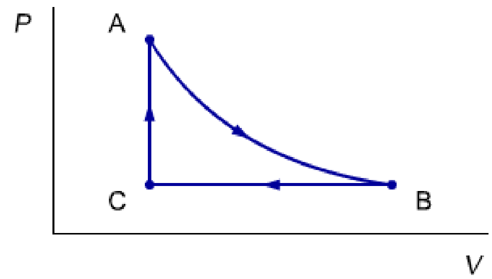
__________ J

A.4 The pressure-volume graph shows an engine cycle with four processes: two isobaric, and two constant-volume. The engine uses 0.800 mol of an ideal monatomic gas as its working substance. For one engine cycle, calculate (a) the heat added to the gas, (b) the heat removed from the gas (stated as a positive number), and (c) the net work done by the engine.



(a) __________ J
(b) __________ J
(c) __________ J

A.5 The pressure-volume graph shows an engine cycle where the working substance is an ideal monatomic gas ($\gamma = 5/3$). The process A-B is adiabatic. At point A, the pressure is 2.70×10^5 Pa and the volume is 0.0180 m^3. At point B, the volume is 0.0450 m^3. During a cycle, calculate (a) the heat added to the gas, (b) the heat removed from the gas (stated as a positive number), and (c) the net work done.

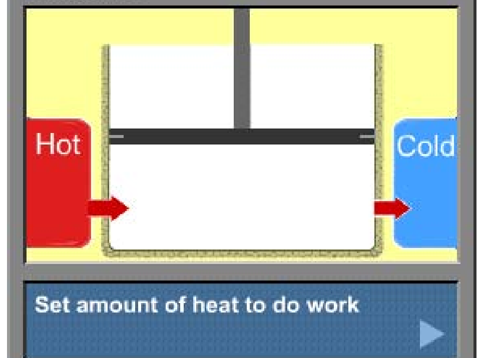


(a) __________ J
(b) __________ J
(c) __________ J

Chapter 22 Second Law of Thermodynamics, Efficiency, and Entropy

22.0 - Introduction

interactive 1



Physics often seems to be about possibilities, but it is also about limits. You can think of the first law of thermodynamics as stating a limit: An engine cannot do more work during an engine cycle than the heat added to it. If it did, it would defy the principle of conservation of energy. If such an engine existed, it would be a source of "free energy." Alas, no such engine exists.

If an engine cannot perform an amount of work greater than the energy added to it, can an engine just "break even"? This is not an idle question: The efficiency of an engine, how much work it does per amount of energy added, is a crucial measure of an engine's utility.

Automobiles, for instance, can take advantage of about one-fourth of the energy produced by the combustion of gasoline to do useful work, making them about 25% efficient.

The second law of thermodynamics establishes the theoretical limit of efficiency: It states that no heat engine can be 100% efficient; that is, no engine can convert all of the heat supplied to it into an equal amount of work. In theory, engines could be built with efficiencies of 99.9%, but never 100%. This limit cannot be reached with any conceivable improvements in engineering; it is a theoretical limit that in principle can never be realized. The second law of thermodynamics and the topic of efficiency are two areas of focus for this chapter.

The concept of entropy provides another way to study thermal processes. In general terms, entropy is a measure of how ordered a system is, and is another useful tool for understanding the efficiency and limits of engines.

The simulation on the right lets you explore the relationship between heat, work and efficiency in an engine cycle. During the engine cycle, the engine will do work and return to its initial condition at the end of the cycle. The heat engine will perform one cycle when you press GO. You can set the amount of heat transferred from the hot reservoir to the engine during the cycle, and the amount of heat the engine expels to the cold reservoir. At the end of the cycle, the work done by the engine and the engine's efficiency are calculated and displayed.

You can add from 50 to 500 joules of heat. These are small amounts of heat, appropriate for a toy or model engine. More than 500 joules will exceed the safety limits of the engine.

We want you to observe two principles at work in this process. First, apply what you learned about the first law of thermodynamics to this engine. What relationship do you expect between the net heat transferred to the engine and the net work done by the engine? (The engine does positive work on the piston as the system expands and raises the piston, and a smaller amount of negative work as the piston falls.) Make a hypothesis and then test it with the simulation.

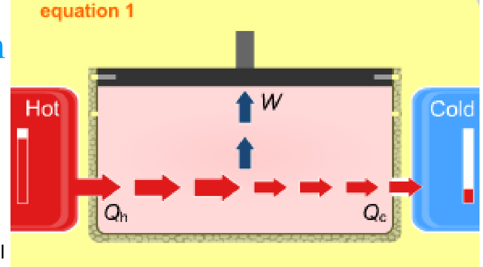
Second, you will encounter the second law of thermodynamics: No engine is 100% efficient. The energy that flows out of the engine in the form of heat to the cold reservoir cannot be used to do useful work. The greater the heat that flows out to the cold reservoir, the less efficient the engine. In the interest of realism, the simulation requires a realistic amount of heat to flow to the cold reservoir.

Even the most efficient practical engines, like those in electric generation plants, run at less than 60% efficiency. See how efficient you can make the simulation engine; it can be much more efficient than the average engine. But, try as you might, you will find that reaching 100% efficiency is a goal you cannot achieve.

22.1 - Efficiency

Efficiency: The ratio of the net work done by an engine during a cycle to the heat energy supplied to the engine.

equation 1



Engine efficiency

$e = W/Q_h$

$e = 1 - \dfrac{Q_c}{Q_h}$

Engine design has been refined for hundreds of years, as engineers have sought to increase the power and reliability of engines while decreasing their size. Equally importantly, they have sought to increase the efficiency of engines, the amount of useful work an engine does divided by the amount of thermal energy supplied to it.

The first formula in Equation 1 states the definition of the efficiency of a heat engine. The efficiency equals the net work W done by the engine divided by the heat Q_h transferred to the system from the hot reservoir. This ratio is often stated as a percentage. A typical internal combustion automobile engine has an efficiency of about 25%, while the diesel and coal powered engines in electrical plants have efficiencies ranging from 40% to 60%. These numbers are for the engines alone. The overall systems − the entire car, the whole electrical plant − run at lower total efficiencies due to inefficiencies outside the engines.

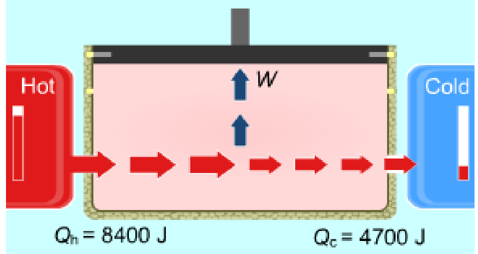
The second formula for engine efficiency shown in Equation 1 provides a way to calculate engine efficiency from the heat Q_h added to the engine, and the heat Q_c that flows out.

The second equation for engine efficiency is derived below from the definition (the first equation). The derivation uses the first law of thermodynamics: The net heat flow equals the work done by the engine plus its change in internal energy. Since the internal energy of an engine is the same at the beginning and end of a cycle, the net heat flow equals the work done by the engine.

Step	Reason
1. $e = W/Q_h$	definition of efficiency
2. $W = Q_h - Q_c$	first law of thermodynamics
3. $e = \dfrac{Q_h - Q_c}{Q_h}$	substitute equation 2 into equation 1
4. $e = 1 - \dfrac{Q_c}{Q_h}$	divide

e = engine efficiency
W = net work during engine cycle
Q_h = heat in during cycle
Q_c = heat out during cycle

example 1



What is the efficiency of this engine?

$e = 1 - \dfrac{Q_c}{Q_h}$

$e = 1 - (4700\text{ J})/(8400\text{ J})$

$e = 0.44 = 44\%$

22.2 - Second law of thermodynamics

Second law of thermodynamics: No heat engine can transform 100% of the energy supplied to it into work during a cycle.

There are several equivalent ways to express the second law of thermodynamics. The definition above states the law in a form that is one of its important consequences: There is a limit to the amount of work that can be done by a heat engine supplied with a certain amount of energy during a cycle. This is called the *Kelvin-Planck* statement of the second law.

To take a step back for a moment: The first law of thermodynamics states that the net heat transferred to a system equals the net work done by the system plus the change in its internal energy. That provides one limit to how much work an engine can do: no more than the net heat transferred to it. In essence, the first law is about energy conservation. Energy cannot be created by a process; it must stay constant.

The second law is a little more dire: It says that during an engine cycle, engines do **less** work than the energy transferred into them. In principle, the work could equal 99.9% of the energy that flows into an engine, but it can never equal 100%.

This law is not about designing an efficient engine. It is instead a physical law that limits the efficiency of any engine. No engine can be 100% efficient.

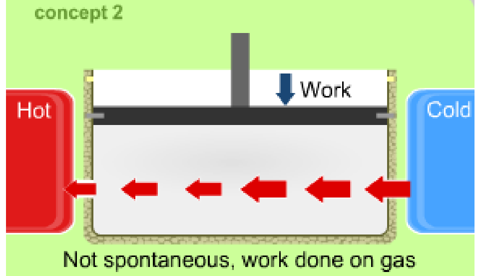
Another way to state the second law, called the *Clausius statement*, is: There can be no process whose sole final result is the transfer of heat from a cooler object to a warmer object. Heat flows spontaneously from an object at a higher temperature to an object at a lower temperature. Heat will **not** flow in the other direction unless compelled to do so. This direction of heat flow agrees with your everyday experience. If you place a quart of ice cream in a hot car, you expect the ice cream to get warmer, not colder. Energy flows from the hotter air to the cooler ice cream.

Stating the law in this fashion may help you to understand why no engine can be 100% efficient. The illustrations at the right show a conceptual diagram of a heat engine. Heat is allowed to flow into the engine from the hot reservoir. Some heat also flows out to the cold reservoir. The maximum amount of work the engine can do during a cycle is the net flow of the heat, the heat in minus the heat out.

concept 1

Second law of thermodynamics
Engines cannot transform 100% of heat into work during a cycle
Heat flows spontaneously only from hot to cold

concept 2



Non-spontaneous heat flow
Work required to force heat to flow from cold to hot

An enterprising engineer might think she could “recycle” the heat absorbed from the engine by the cold reservoir, say by connecting the cold reservoir to the hot reservoir. The Clausius statement of the second law says that heat will not flow spontaneously from the cold reservoir to the hot.

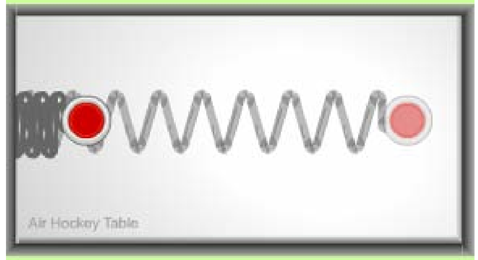
To solve this problem, she might attach another engine to force heat to flow from the cold reservoir to the hot. This is indeed possible: Air conditioners and refrigerators cause heat to flow from a cooler region to a warmer one. However, this heat flow is not spontaneous or free; it requires work and comes at a price, as the electrical bills for air conditioners and refrigerators indicate.

22.3 - Reversible and irreversible processes

Reversible process: A process in which a system can be returned to its initial state without the addition of energy.

Irreversible process: Energy must be added to a system to return it to its initial state after an irreversible process occurs.

concept 1



Reversible process
System can be returned to initial state without adding energy

Some of the crucial underpinnings of the theories concerning engine efficiency rely on the concepts of reversible and irreversible processes. In general, a process is a series of steps that move a system from one state to another. When a system undergoes a reversible process, it can be returned to its initial state without the addition of energy.

Every real process is irreversible, although some are close to being reversible. Consider, for example, a puck attached to a spring on an air hockey table, as shown in Concept 1. Imagine the spring is initially compressed and the puck is held in place. When the puck is released, the compressed spring will push the puck to the right until it pauses momentarily. This is a process: The spring pushing the puck until it pauses. In an ideal system (no friction, no air resistance), the puck will return to its initial position as the spring contracts with no additional energy required. The process of spring expansion can be thought of as reversible.

concept 2

Irreversible process
Cannot be reversed without adding energy
Video runs only "one way"

In contrast, if you drop an egg and it breaks, several changes take place that cannot be easily reversed (that is quite an understatement). For example, the breaking of the egg creates sound energy through vibrations in the air. This reduces the energy of the egg. You do not expect to be able to put it back together again without extraordinary effort. This is an irreversible process. It would take an extreme amount of time and effort (Humpty Dumpty inevitably comes to mind) to return the egg to its initial state.

One way to think about reversibility is to imagine videotaping a process. If you can easily decide whether the tape is being played forward or backward, the process is irreversible. If you watched a videotape of an egg shattering, you would know the direction the tape is being played. On the other hand, with the puck and the spring, you would not be able to discern easily if the tape were being played forward or backwards. (If you watched it for several cycles of motion, however, you would note that the extreme positions of the puck were less far from equilibrium as the system lost energy to dissipative forces like air resistance and friction.)

These principles can be stated in terms of a system and its environment. With an irreversible process, a system can be returned to its initial state but its environment must change. This is particularly applicable to heat engines, where the “environment” may be modeled as hot and cold reservoirs. The engine mechanism, the gas and the piston, can be returned to their initial state, but their environment changes. The hot reservoir becomes a little less hot and the cold reservoir becomes warmer each time the engine completes a cycle.

Reversibility provides a conceptual tool for dealing with thermodynamics. The thermodynamic processes we have studied − constant-volume, isobaric, adiabatic, and isothermal − can be viewed as reversible under the appropriate conditions. For example, imagine transferring thermal energy to a quantity of gas so as to expand it **very** slowly, so that the temperature of the gas does not change. This is an isothermal expansion that would travel a particular path on a pressure-volume graph. To reverse the process, we could let the gas expel heat equally slowly, causing it to contract in the opposite direction along the same path. The environment (hot or cold reservoir) with which the gas exchanges heat must have a temperature infinitesimally different from that of the gas, so that the process proceeds slowly enough. This makes for a very efficient process, since no “waste” heat is created. However, because the process is extremely slow, it is not very useful in most practical applications.

22.4 - Free expansion

Free expansion: A gas expands irreversibly into a vacuum.

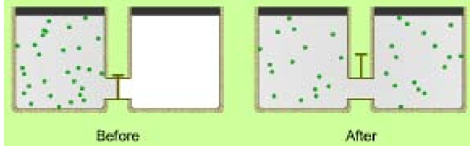
Free expansion supplies an example of an irreversible process. Above, you see an illustration of free expansion. When the valve is opened, the gas on the left rushes in to fill the vacuum on the right. This process occurs rapidly and chaotically, and during the process, properties such as pressure are not well defined for the system, since they vary by location. This is an example of an irreversible process.

Another way to characterize an irreversible process is that the system passes through non-equilibrium states; these are unsteady states in which the system would never settle naturally. If you examine a snapshot of a non-equilibrium state, you know that the system is just passing through the state, and will not naturally pause or stop there. An example of a non-equilibrium state for the free expansion process is the situation just after the valve has been opened, while much of the gas is in the left chamber. You would be very surprised to see the system remain in that non-equilibrium state.

A free expansion does no work. Because a vacuum exerts no pressure, no force opposes the expansion of the gas, which means no work is done. The process is adiabatic because the walls of the chamber are insulated, and because the process occurs too rapidly for heat transfer to occur. However, this is not a reversible adiabatic process, because it occurs quickly and chaotically. The pressure and volume are not well defined at each instant, so the product PV^{γ} cannot be considered to be constant, as it would be for a reversible adiabatic expansion of an ideal gas.

Since neither heat transfer nor work occurs, there is no change in the internal energy of the gas, so its temperature remains constant. That is, the free expansion is also isothermal. These statements are expressed as a set of equations on the right.

concept 1



Free expansion
Gas quickly expands into vacuum
Chaotic: pressure and volume not well-defined during expansion

equation 1

Free expansion

$W = 0$

$Q = 0$ (adiabatic, but $PV^{\gamma} \neq$ constant)

$\Delta E_{\text{int}} = 0$ (isothermal)

22.5 - Entropy

Entropy: A property of a system. When heat is transferred to a system, its entropy increases.

Entropy is a concept used to describe the state of a system, and it is a property of a system, just like pressure, volume, temperature or internal energy. Some properties may be directly measured or read from a gauge. A thermometer will tell you an object's temperature and a meter stick will tell you its length. For other properties, such as kinetic energy, there are no direct measures; there are no "KE gauges." A property like kinetic energy must be computed from factors that can be measured, mass and speed.

Similarly, there are no direct "entropy gauges" available to scientists or students. In this section, we focus on the relationship of entropy, temperature and heat without concerning ourselves too much about exactly what is meant by "entropy." In other words, we will start our discussion of entropy by considering some properties that can be observed and are familiar to you, and focus on entropy itself once this groundwork has been laid. In fact, the concept of entropy was historically developed in a process like this. It arose out of the quest to understand the relationship between temperature and heat.

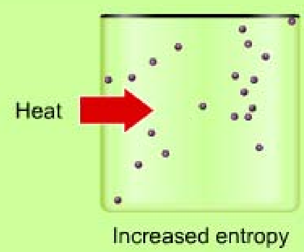
When heat flows into a system, its entropy increases. When heat flows out of a system, its entropy decreases. The change in entropy equals the heat divided by the temperature at which the heat flow occurs.

We express this as a formula in Equation 1. To apply this equation, the process must be reversible and the temperature measured in kelvins. Since absorbing or expelling heat will change the temperature of the system, this change in entropy must be measured for a small amount of heat, or the system must be large enough that it can expel or absorb a fair amount of heat with only a negligible change in temperature. The units for entropy are joules/kelvin.

This equation is useful for two reasons. First, it provides the tool for computing a change in entropy. Second, it (finally!) provides a definition of temperature more formal than just "something measured by a thermometer": Temperature is the slope of an entropy-heat curve. (Solving the equation on the right for temperature shows why this is the case.)

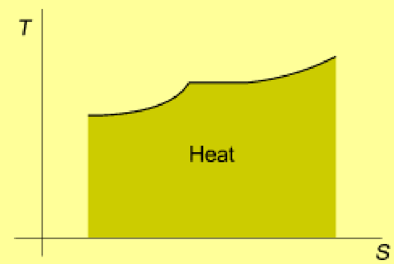
The entropy of a system, like any property, depends only on the system's state, not how the system arrived there. There are many processes that change a system from a particular initial state to a particular final state. The system's resulting change in entropy is the same for any of these processes. (This is analogous to gravitational potential energy, where only the initial and final positions of an object matter, not the path it took

concept 1



Entropy
Property of objects, systems
Increases as heat is transferred to object

equation 1



Entropy

$$\Delta S = \frac{Q_{\text{rev}}}{T}$$

ΔS = change in entropy
Q_{rev} = net heat transferred in reversible process

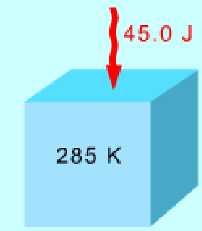
between them.)

At the right, you see a graph of temperature plotted against entropy. As heat flows into an object, its temperature increases, as does its entropy. The area under the curve, positive or negative, equals the amount of heat transferred.

The curve is not a straight line. The horizontal section occurs at a first-order phase transition, where heat added causes the substance to change from solid to liquid, or liquid to gas. At phase transitions, adding heat increases the entropy, but not the temperature.

T = temperature (K)
Units: joules/kelvin (J/K)

example 1



Heat flows to a large block in a reversible process. What is the change in the entropy of the block? Assume its temperature is constant.

$$\Delta S = \frac{Q_{rev}}{T}$$

ΔS = (45.0 J)/(285 K)
ΔS = 0.158 J/K

22.6 - Interactive checkpoint: ice cube entropy

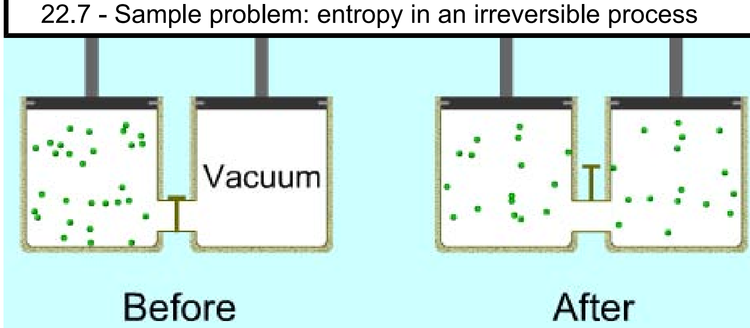
Calculate the change in entropy as an 8.25-gram ice cube melts in a reversible process at 0°C. Then calculate the change in entropy as the same amount of water is vaporized in a reversible process at 100°C. (Do not include the change in entropy as the water is heated from 0°C to 100°C.)

Answer:

ΔS_m = ________ J/K

ΔS_v = ________ J/K

22.7 - Sample problem: entropy in an irreversible process



In a free expansion, 1.50 moles of nitrogen doubles in volume. What is the change in entropy? Assume the nitrogen behaves like an ideal gas.

This is a challenging problem that shows how the entropy change resulting from a process in a system can be calculated, even if the process, like the free expansion shown above, is not reversible.

The equation for the change in entropy cannot be applied to this process because it is irreversible. However, there is a way to solve the problem, using an important principle. The change of entropy does **not** depend on the process; the change is always the same between any two states, however the change occurred.

This means that to solve the problem, we can use a different process, one which is reversible, which leads to the gas being at the same final state as the free expansion process. By finding a reversible process that takes the gas from the same initial state to the same final state as the free expansion, we can calculate the change in entropy using that process instead. We need a reversible process that leads to the same volume of gas, and a process during which the temperature does not change, since free expansion occurs at a constant temperature.

Our substitute process is a reversible isothermal expansion, in which the gas volume changes with no change in temperature.

Variables

temperature	T
initial volume	V_i
final volume	$V_f = 2V_i$
amount of gas	$n = 1.50$ mol
gas constant	$R = 8.31$ J/mol·K
heat added	Q_{rev}

What is the strategy?

1. Find a reversible process that takes the system between the same initial and final conditions. We use an isothermal process. Calculate the heat added for this process.
2. Apply the entropy equation to calculate the change in entropy.

Physics principles and equations

The entropy equation

$$\Delta S = \frac{Q_{rev}}{T}$$

For an isothermal process, the work equals the heat transferred, and the work can be calculated from the initial and final gas volumes.

$$W = Q$$

$$W = nRT \ln(V_f/V_i)$$

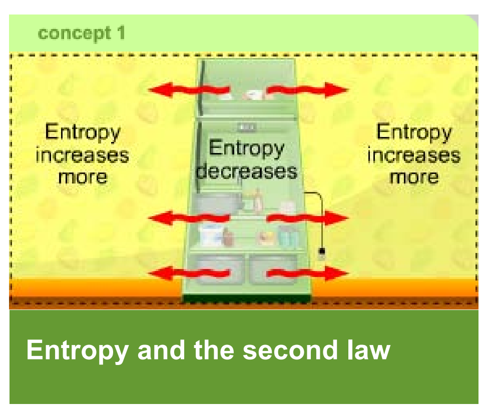
Combining these two equations yields an equation for the heat associated with the reversible process.

$$Q = nRT \ln(V_f/V_i)$$

Step-by-step solution

Step	Reason
1. $\Delta S = Q_{rev}/T$	entropy equation
2. $Q_{rev} = nRT \ln(V_f/V_i)$	isothermal process
3. $\Delta S = nRT \ln(V_f/V_i)/T$	substitute equation 2 into equation 1
4. $\Delta S = nR \ln(2)$	simplify
5. $\Delta S = (1.50 \text{ mol})(8.31 \text{ J/mol·K})\ln(2)$	substitute values
6. $\Delta S = 8.64$ J/K	evaluate

22.8 - Second law of thermodynamics: entropy



Second law of thermodynamics: Entropy increases, or at best remains constant, in any isolated system.

If nature abhors a vacuum, as the saying goes, then it revels in messes. The universe tends toward increasing disorder. So if your life feels like it is becoming increasingly chaotic, well, you are just going with the flow.

A common non-physicist's statement of the second law is "systems tend toward disorder," a paraphrase of the definition above, which is the second law stated in terms of entropy. Entropy is considered as a measure of a system's disorder. The disorder of a system either stays constant, or increases. Your room never spontaneously becomes more orderly.

You might think that the second law is violated because you can reduce the entropy of an object by cooling it. For instance, you can place a hot drink inside a refrigerator and cool the drink.

However, the second law applies to isolated systems and the refrigerator does not function as an isolated system. If you place your hand by the back of a refrigerator, you will realize that refrigerators emit heat. This heat increases the entropy of the surrounding atmosphere. When we correctly apply the second law to the isolated system of the refrigerator and surrounding atmosphere, we find that entropy stays constant, or increases.

A series of reversible processes can, in theory, leave the entropy of a system unchanged. In practice, however, no process is perfectly reversible, and the entropy of an isolated system increases as processes occur.

Above, we asserted that the heat flow out of a refrigerator would increase the entropy of a system. We will show how this is true, using a particular instance.

We rely on the formulation of the second law that states that heat flows spontaneously only from a hotter object to a colder one. Also recall that the entropy change equals the heat flow divided by the temperature at which the heat flow occurs. The colder the temperature, the larger the increase in entropy.

In Example 1 on the right, 1100 J of heat flows from an object at 250 K to an object at 130 K. To calculate the system's entropy change, we consider the two blocks separately. For the warmer object, we divide the heat by the temperature of the warmer object. Since heat flows out of the warmer object, that heat flow is negative. The change in entropy of the hotter object is −4.4 J/K. Since heat flows into the cooler object, that heat flow is positive. Doing similar calculations for the cooler object, but dividing by the cooler object's lower temperature, we determine that its entropy change is +8.5 J/K. The net change in entropy for the system is 4.1 J/K. The positive sign for the change in entropy tells us that the entropy increases and the second law is obeyed.

This example helps to illustrate why entropy increases with spontaneous heat flow. Heat flows from hot to cold, and as the example illustrates, the reduction of entropy in the hotter object is more than matched by the increase in entropy of the cooler object. This must always be the case since the change in entropy equals the heat divided by the temperature, so a hotter object "expels" less entropy than the colder object "absorbs." In our example, we treated the two objects as being at constant temperatures, but even if their temperatures changed, the spontaneous flow of heat would still cause the entropy of the system to increase.

The entropy of an isolated system can only increase
The total entropy of the universe is increasing
Carpe diem

equation 1

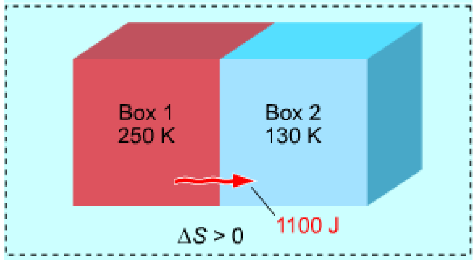
Entropy and the second law

In an isolated system:

$$\Delta S \geq 0$$

S = entropy

example 1



What is the change in entropy of the system during this spontaneous heat transfer?

$$\Delta S = \Delta S_1 + \Delta S_2$$

$$\Delta S = \frac{Q_1}{T_1} + \frac{Q_2}{T_2}$$

$$\Delta S = \frac{-1100\ \text{J}}{250\ \text{K}} + \frac{1100\ \text{J}}{130\ \text{K}}$$

$$\Delta S = -4.4 + 8.5 = 4.1\ \text{J/K}$$

22.9 - Entropy and disorder (oops: disorder)

Entropy: Sometimes defined as a measure of the disorder of a system.

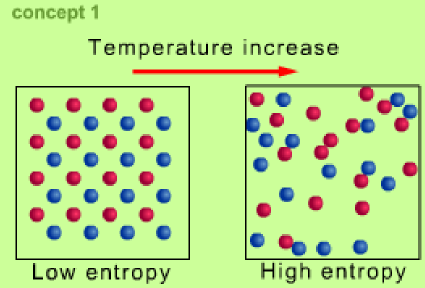
The concept of entropy as disorder has entered popular culture thoroughly enough that this section both uses, and critiques, this popular definition. Entropy often is described as measuring the order, or really the disorder of a system. A system that has more disorder has more entropy. For instance, when a new deck of playing cards is shuffled, it goes from being well ordered (the cards are arranged by suit and number) to disordered (the cards can be in many possible arrangements). The deck's entropy has increased.

In thermodynamics, entropy is more often discussed in terms of heat and temperature. The entropy of an object increases as its temperature increases. For instance, when ice is heated so that it melts, and then the liquid is further heated until it turns into steam, the water's entropy increases. On the other hand, in ice the water molecules are "ordered" because of their crystalline structure. In a gas, the molecules have no defined positions, they move randomly, and a particular water molecule may wander far from its initial position. Transferring thermal energy into the system increases its temperature, and its "disorder" increases as its temperature increases.

In some respects these examples work well, providing a visual sense or metaphor to the abstraction of entropy. On the other hand, they are not particularly rigorous, and they do not necessarily link the underlying reality of entropy − its relation to heat − to what we actually see.

For instance, what we perceive as more "ordered" might in fact have greater entropy. Consider the example of ice cubes and liquid water. What looks more "orderly": a glass containing jagged irregular chunks of ice, or the same glass filled with an equivalent amount of liquid water? The glass of water "looks" more orderly, but considering water at the molecular level, the ice is more orderly.

concept 1



Entropy and disorder

"Popular" definition of entropy: measure of disorder
As temperature increases, so does disorder

Something is also missing if entropy is only considered to be the "disorder" of a system. Consider an unusual trick deck of cards in which every card is the two of spades. You can shuffle all you like without creating a more "disordered" deck. Entropy as disorder depends on the number of different ways the elements of a system can be arranged. For instance, you can create more disorder by shuffling a standard 52-card deck than one that consists of, say, only 13 cards since the 52-card deck can be arranged in many more ways (8.07×10^{67}) than can the 13-card one (only 6.23×10^{9}).

In summary, considering entropy as disorder, and increasing entropy as increasing disorder, can be a useful metaphor at times, but it also has its limits and potential traps.

22.10 - Maximum engine efficiency and reservoir temperatures

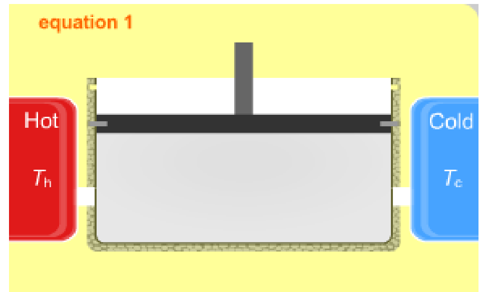
Heat engines are everywhere in the modern world. They convert heat into useful work. These engines always consist of a "hot" source reservoir from which thermal energy is taken − for example, the coal burning firebox in an old-fashioned train engine − and a "cold" exhaust reservoir where thermal energy is expelled. In the case of a train's steam engine, that cold reservoir is the atmosphere.

Higher efficiency − turning more of the heat into work − is what heat engine designers strive for. This results in less fuel consumption, which saves money, extends natural resources and reduces pollution.

We can state an inequality, shown on the right, expressing the maximum efficiency of any engine as determined by the temperatures of its hot and cold reservoirs. In this section, we derive this inequality. To do so, we consider the engine, including its reservoirs, as an isolated system, and assume the reservoirs are large enough that the heat flow does not appreciably change their temperatures.

During an engine cycle, the entropy of the hot reservoir decreases as heat flows out of it and the entropy of the cold reservoir increases as heat flows into it. The entropy of the gas does not change after the completion of a cycle because the engine returns to its initial state.

The second law of thermodynamics dictates that the entropy of an isolated system must increase or stay the same during any process, including a complete engine cycle. This means the decrease in entropy of the hot reservoir must be matched or exceeded by the increase in entropy of the cold reservoir. The net entropy increases.



Maximum efficiency of heat engine

$$e \le 1 - \frac{T_c}{T_h}$$

e = efficiency
T_c = temperature of cold reservoir (K)
T_h = temperature of hot reservoir (K)

Variables

engine efficiency	e	
net change in entropy	ΔS	
	hot reservoir	cold reservoir
temperature	T_h	T_c
heat transferred	Q_h	Q_c
change in entropy	ΔS_h	ΔS_c

What is the strategy?

1. Calculate the change in entropy for the hot and cold reservoirs. The net change in entropy, which cannot be negative, is the sum of these two quantities. Write this as an inequality, and then rearrange the inequality so the ratio of reservoir temperatures is on one side.
2. Apply the definition of engine efficiency to the inequality.

Physics principles and equations

The definition of entropy

$$\Delta S = \frac{Q_{rev}}{T}$$

The entropy statement of the second law

$$\Delta S \ge 0$$

An equation for engine efficiency

$$e = 1 - \frac{Q_c}{Q_h}$$

Step-by-step derivation

We start with the entropy inequality, and use that to write an inequality with the ratio of reservoir temperatures on one side.

Step	Reason
1. $\Delta S \geq 0$ $\Delta S_c + \Delta S_h \geq 0$	entropy never decreases
2. $\Delta S_c = \frac{Q_c}{T_c}$	definition of entropy, cold reservoir
3. $\Delta S_h = -\frac{Q_h}{T_h}$	definition of entropy, hot reservoir
4. $\frac{Q_c}{T_c} - \frac{Q_h}{T_h} \geq 0$	substitute equations 2 and 3 into inequality 1
5. $\frac{Q_c}{Q_h} - \frac{T_c}{T_h} \geq 0$	multiply by T_c/Q_h

The inequality in step 4 is a concise statement of the entropy changes during an engine cycle: The increase in entropy of the cold reservoir outweighs the decrease in entropy of the hot reservoir.

Now we use the equation for engine efficiency to prove the inequality that describes the maximum engine efficiency.

Step	Reason
6. $e = 1 - \frac{Q_c}{Q_h}$	equation stated above
7. $e \leq 1 - \frac{Q_c}{Q_h} + \left[\frac{Q_c}{Q_h} - \frac{T_c}{T_h}\right]$	add inequality 5 to equation 6
8. $e \leq 1 - \frac{T_c}{T_h}$	Simplify

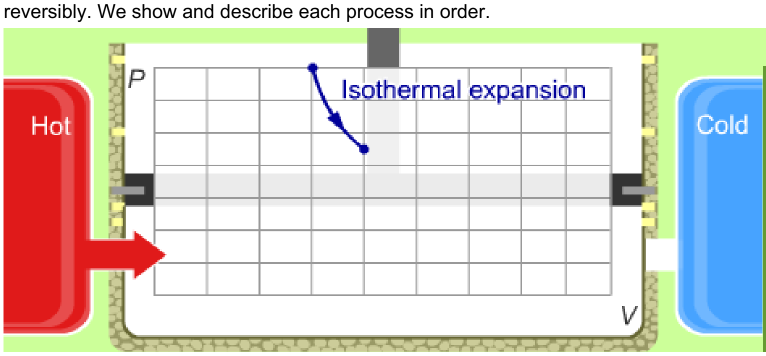
The inequality derived in step 8 shows the maximum possible efficiency for an engine with given hot and cold reservoir temperatures. The efficiency limit $1 - T_c/T_h$ is maximized when the cold reservoir is as cold as possible, and the hot reservoir is as hot as possible. For instance, an engine that has a cold reservoir of 300 K and a hot reservoir of 1000 K has an efficiency limit of 1 − 0.3, or 70%. One appeal of diesel engines is that they run "hotter" than gasoline engines, which is a reason why diesel engines are more fuel-efficient.

There are practical limits to engine efficiency. Reservoirs at 10 K and 10,000 K could make for an extremely efficient engine in principle, but it is not possible to build a cost-efficient engine that would maintain reservoirs at these temperatures. Factors other than efficiency, such as cost and the ability to supply power, also affect the design of engines.

22.11 - Carnot cycle

In 1824, the French scientist Sadi Carnot devised the theoretical workings of an engine cycle now known as the *Carnot engine*. The engine used four reversible processes. Carnot stated that an engine that uses entirely reversible cycles would be maximally efficient.

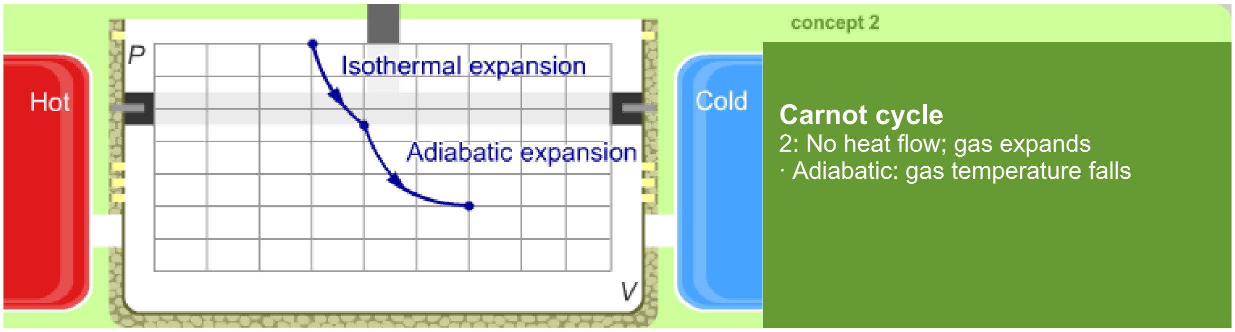
Before delving further into the cycle's efficiency, we first focus on the four processes that make up the cycle. All these processes operate reversibly. We show and describe each process in order.



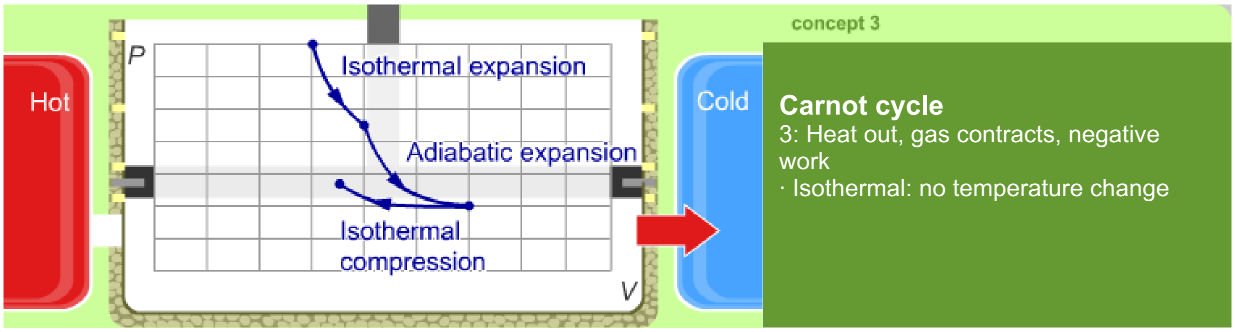
concept 1

Carnot cycle
1: Heat in, gas expands, positive work
· Isothermal: no temperature change

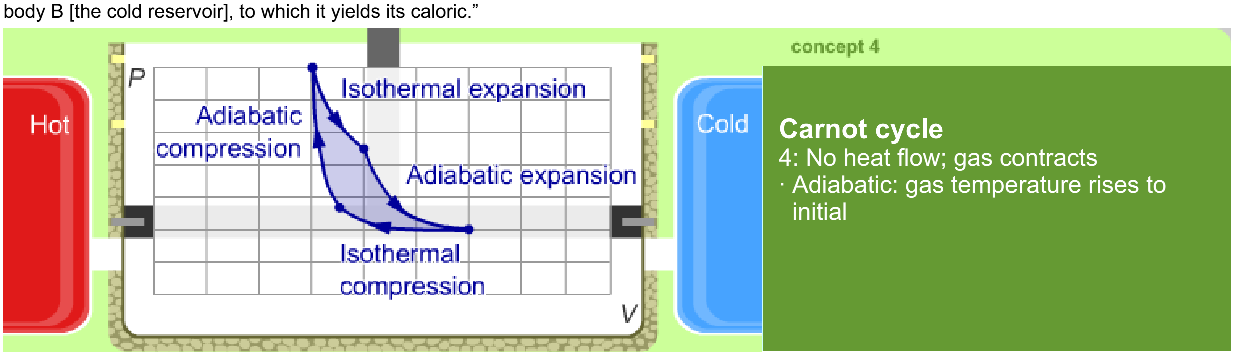
First: As shown directly above, heat is allowed to flow into the engine from the hot reservoir. The gas expands, doing positive work. This process is controlled so that the gas expands at a constant temperature (isothermal expansion). As Carnot wrote, "The body A [the hot reservoir] furnishes the caloric [heat] necessary to keep the temperature constant." The heat flow supplies the energy for the gas to expand.



Second: Heat is no longer allowed to flow into the engine; this process is adiabatic. The gas continues to expand. The energy for this expansion comes from the internal energy of the gas, decreasing its temperature. Carnot stated: "The air [the working gas] is then no longer in contact with any body capable of furnishing it with caloric. The piston meanwhile continues to move."



Third: Heat is now allowed to flow out of the engine into the cold reservoir. The piston lowers and the gas contracts. This means the gas does negative work on the piston. Because the net heat flow equals the work done by the gas, the gas's internal energy and temperature remain constant. In other words, this is an isothermal contraction step. Carnot: "The air remains at a constant temperature because of its contact with body B [the cold reservoir], to which it yields its caloric."



Fourth: Again, the flow of heat is stopped. The gas continues to contract as the piston moves down, meaning the system (that is, the gas) does negative work. The negative work of the gas increases its internal energy and temperature. This is an adiabatic contraction cycle. Carnot: "The compression of the air is continued, which being then isolated, its temperature rises." The gas is returned to its initial temperature and the piston returns to its initial position. The cycle is complete.

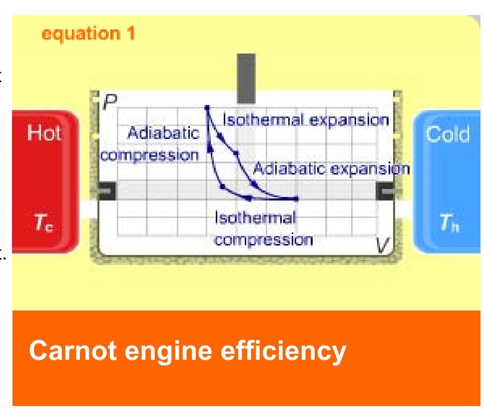
The amount of work done by the engine equals the area enclosed by the graphs of the four processes. Although of considerable conceptual interest, this engine cycle cannot be used in practical applications. Why? Since the processes must operate extremely slowly so that they are reversible, an engine based on this cycle would generate little power.

22.12 - Carnot cycle and efficiency

Carnot proved that any fully reversible engine, like his, was the most efficient possible. In his argument, this French scientist showed that a more efficient engine cycle would violate the second law of thermodynamics, which meant it was not possible to construct such an engine. In an irreversible process, energy is lost from the system to its environment in unrecoverable ways, such as through friction or sound energy. All real engines operate irreversibly and are less efficient than a Carnot engine.

Earlier, we showed that the maximum efficiency of a heat engine was **less than or equal** to one minus the ratio of the temperature of the reservoirs. The closer an engine cycle comes to having no increase in entropy, the closer its efficiency will be to this limit.

The Carnot cycle attains the theoretical maximum efficiency for any engine functioning with reservoirs at two particular temperatures. It can be shown that the efficiency of this cycle **equals** one minus the ratio of the temperature of the cold reservoir to the hot reservoir. This important conclusion is shown in Equation 1.



Even the ideal Carnot engine is not 100% efficient. To make a Carnot engine operate at 100% efficiency, its cold reservoir would have to be at absolute zero (which is theoretically impossible), or its hot reservoir would have to be infinitely hot.

$$e_{ce} = 1 - \frac{T_c}{T_h}$$

e_{ce} = efficiency of Carnot engine
T_c = temperature of cold reservoir (K)
T_h = temperature of hot reservoir (K)

22.13 - Otto cycle: internal combustion engine

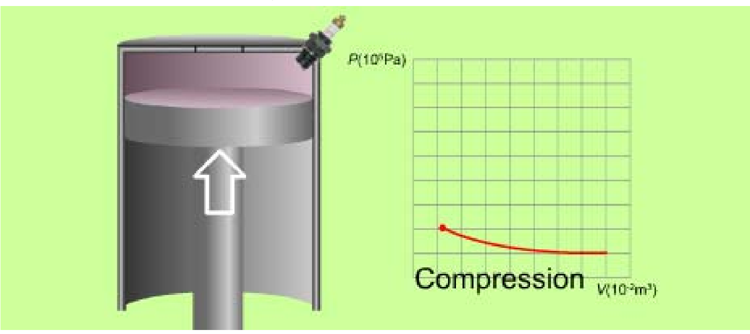
The internal combustion engine shown above is used in automobiles and is probably familiar to you. In this section, we describe the operation and engine cycle of this engine, and present an equation for its efficiency. The cycle is called the Otto cycle, after the German inventor of the "four-stroke" internal combustion engine, Nikolaus August Otto. A "stroke" is a movement in or out of a piston. In the complete Otto cycle, a piston moves in and out twice, making four strokes.

Internal combustion engine.

In an internal combustion engine, there is no heat reservoir in the usual sense. Instead of heat flowing into the engine, the working gas itself is the source of energy. It is a mixture of air and gasoline vapor that is ignited inside a cylinder (hence the term "internal combustion"), and expands to drive a piston that does work. The thermal energy that is used to do work is created by burning the gasoline mixture. You see the cylinder and piston in the illustrations below.

The four steps of the basic thermodynamic cycle together require two strokes of the piston. The steps occur rapidly. Two additional strokes drive the combustion products out of the cylinder and replace the air and gasoline mixture for the next cycle. For simplicity, we combine these two strokes into one "exhaust" step.

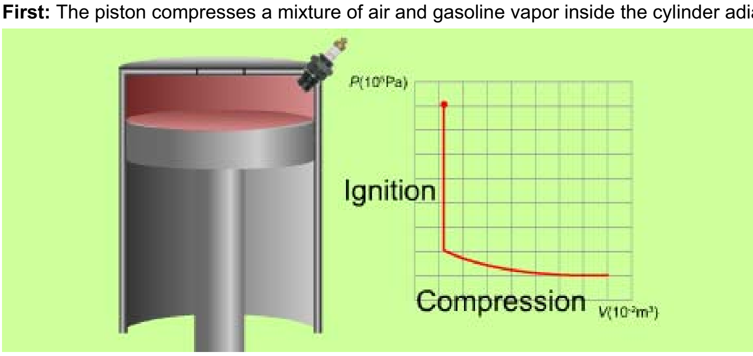
Here are the steps in the Otto cycle.



concept 1

Otto cycle (internal combustion)
· Air-fuel compressed adiabatically

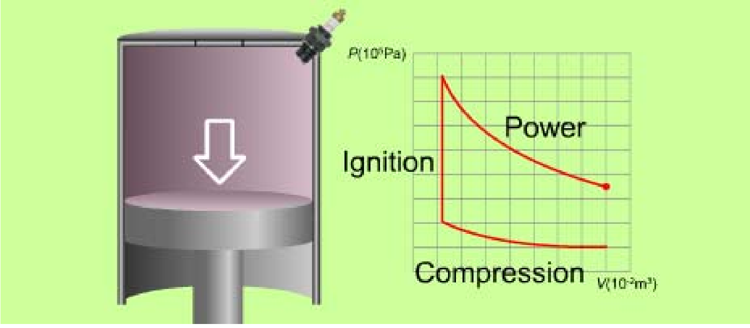
First: The piston compresses a mixture of air and gasoline vapor inside the cylinder adiabatically.



concept 2

Otto cycle (internal combustion)
· Ignition: temperature and pressure increase

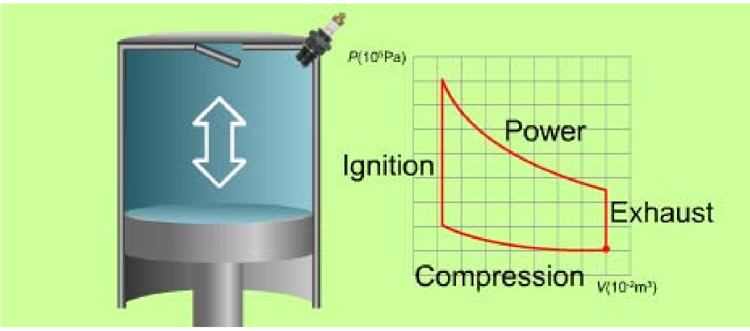
Second: A spark plug ignites the mixture. This is a constant-volume process: The temperature and pressure of the mixture both increase, but the volume stays the same.



concept 3

Otto cycle (internal combustion)
· Mixture expands, doing work

Third: The gas expands adiabatically, moving the piston to do work.



concept 4

Otto cycle (internal combustion)
· “Exhaust” expelled, new air/gasoline injected

Fourth: Finally, there is an “exhaust” step involving both an up and a down stroke of the piston where the combusted mixture in the cylinder is replaced through some valves by a fresh mixture of air and gasoline at lower pressure and temperature. The piston moves in and out, but it returns to its original position so this is considered a constant-volume process, as shown on the pressure-volume graph. The cycle then begins again.

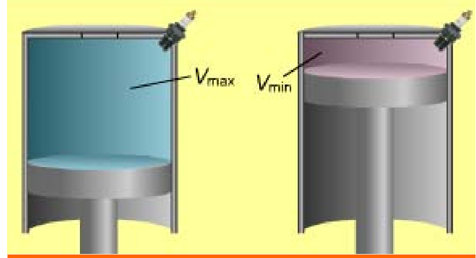
The theoretical efficiency of the engine can be computed as shown in Equation 1 on the right. V_{max} is the volume of the cylinder when the piston is “out”. That is the maximum volume, and V_{min} is the minimum volume. The ratio of the molar specific heats for the air-fuel mixture C_P/C_V is indicated by γ (the Greek letter gamma). The ratio V_{max}/V_{min} is called the *compression ratio* of the engine. The higher the compression ratio, the more efficient the engine.

In Example 1 on the right, you see an example problem that computes the theoretical efficiency of a typical car engine. The actual efficiency of real car engines ranges from 20-26%, less than the equation predicts, due to the heat expelled from the engine and other factors.

You may be familiar with *octane ratings* for gasoline. Typical octane ratings, as seen on the pumps at gas stations, are in the range of 85 to 95. The octane rating of a gasoline indicates the tendency of the gasoline-air mixture to self ignite when it is compressed. The higher the octane rating, the more compression the mixture can withstand before combusting spontaneously. High compression gasoline engines require high octane fuel to prevent “knocking” that occurs when the fuel mixture in the cylinder ignites before the compression cycle is complete, creating a shockwave that collides with the piston.

A diesel engine operates in a similar fashion to a gasoline engine, but does not require a spark plug. Instead, air alone is compressed in the cylinder, causing it to reach very a high temperature. Then, fuel is injected into the cylinder, where it is ignited due to the high temperature. To create the high temperature needed to ignite the fuel, diesel engines have higher compression ratios than typical gasoline engines. This means they are more efficient than gasoline engines.

equation 1

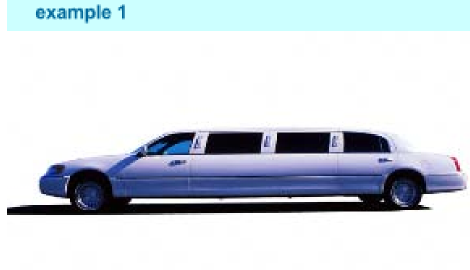


Otto cycle efficiency

$$e_{ice} = 1 - \frac{1}{(V_{max}/V_{min})^{\gamma - 1}}$$

e_{ice} = efficiency of internal combustion engine
V_{max} = max volume of cylinder
V_{min} = min volume of cylinder
V_{max}/V_{min} is "compression ratio"
$\gamma = C_P/C_V$, ratio of molar specific heats

example 1



The car engine has a compression ratio of 8.0. The

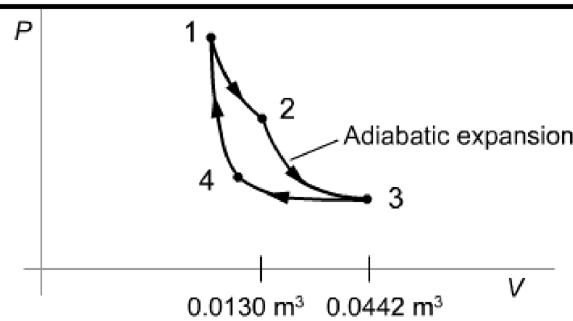
molar specific heat ratio γ is 1.4. What is the theoretical engine efficiency?

$$e_{ice} = 1 - \frac{1}{(V_{max}/V_{min})^{\gamma - 1}}$$

$$e_{ice} = 1 - \frac{1}{(8.0)^{1.4-1}}$$

$$e_{ice} = 0.56 = 56\%$$

22.14 - Interactive checkpoint: Carnot cycle efficiency



The graph shows the cycle of a Carnot engine that uses nitrogen gas. What is the efficiency of this engine?

Answer:

$e =$ ______

22.15 - Heat pumps

Heat pump: A device that transfers heat from a colder environment to a warmer one.

It may sound a bit odd to speak of transferring heat from a colder environment (say the outdoors in the winter) to a warmer one (the interior of a building), but that is what heat pumps do: They pump heat opposite to the direction it would normally flow.

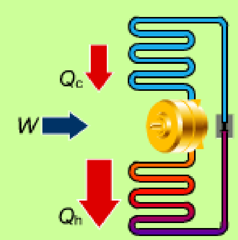
Because heat pumps are more energy-efficient than typical furnaces, they are becoming increasingly popular. An additional and very attractive benefit of heat pumps is that in the summertime, they can be run in reverse to pump heat out of the cool house into the warm outdoors, so the same appliance does double duty as a furnace and as an air conditioner.

A heat pump can be thought of as a heat engine run in reverse. With a heat engine, heat flows from the hot reservoir to the engine and then to the cold reservoir, and work is done **by** the engine. With the pump, heat flows the opposite direction, from the cold reservoir to the hot, and work must be done **on** the engine to accomplish this, since heat does not flow spontaneously from cold to hot.

As mentioned above, a heat pump is in some ways analogous to an air conditioner (or a refrigerator). In both cases, work is performed on the device, and it is used to transfer heat in the direction opposite to the direction in which it spontaneously flows. However, there is a difference between the two appliances. A heat pump moves heat from the cold outdoors to a warmer indoors, with the purpose of further increasing the temperature of the indoors. An air conditioner moves heat from indoors to outdoors, with the purpose of further cooling the interior.

Efficiency is analyzed somewhat differently with heat pumps than it is with heat engines.

concept 1



Heat pump
Work moves heat from cold to hot reservoir

equation 1

Coefficient of performance

$$COP = \frac{Q_h}{W}$$

COP = coefficient of performance
Q_h = energy conveyed to hot reservoir

Recall that the maximum efficiency of a heat engine is determined by the ratio of the cold and hot reservoir temperatures ($e = 1 - T_c/T_h$). When the temperature difference is large, the ratio is small and the engine can be more efficient. In contrast, as you can see from Equation 2, heat pumps work best when there is little temperature difference. The greater the difference, the harder the pump has to work to transfer energy from the cold to the hot reservoir. For this reason, heat pumps are rated by the *coefficient of performance* (COP). You see the definition of the coefficient of performance in Equation 1 on the right: It is the heat transferred to the hot reservoir divided by the work done **on** the gas. Both the heat transfer and the work done on the engine are treated as positive quantities.

One way to think of the efficiency of a heat engine is as the ratio of "what you want" (work done) to "what you pay" (heat added). This applies to the COP as well. The COP for a heat pump equals the ratio of "what you want" (heat flow) to "what you pay" (the work done on the gas). The larger the COP, the more heat is transferred to the hot reservoir per unit of work done.

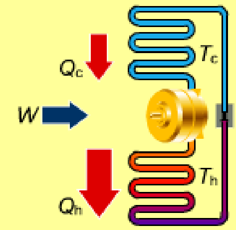
The maximum COP for a heat pump is determined, as for a heat engine, by the temperatures of the hot and cold reservoirs. An inequality that limits the maximum COP is shown in Equation 2 on the right. The right side of the inequality is the ratio of the hot reservoir temperature to the difference of the hot and cold temperatures. A Carnot engine operating "in reverse" as a heat pump achieves the maximum COP. That is, the equality holds.

In Example 1, the maximum COP is calculated for a heat pump operating to warm a house, with an indoor temperature of 21°C (about 70°F) and an outdoor temperature of 13°C (about 55°F). The theoretical maximum COP is 37; the actual operating COP would be less. Typical COPs for heat pumps range from about five to 10. A heat pump with a COP of five would transfer five joules of thermal energy for every joule of work supplied.

A refrigerator acts something like a heat pump, but its purpose is to cool the cold reservoir, the refrigerator's interior. The coefficient of performance of refrigerators is sometimes measured as Q_c/W. Again, this is the ratio of "what you want" (cooling the cold reservoir) to "what you pay" (work done). For refrigerators, COP values from four to six are typical.

W = work done on gas

equation 2



Maximum *COP*

$$COP \le \frac{T_h}{T_h - T_c}$$

T_h = hot reservoir temperature (K)
T_c = cold reservoir temperature (K)

example 1

What is the maximum *COP* of a heat pump operating between these indoor and outdoor temperatures?

$$COP_{max} = \frac{T_h}{T_h - T_c}$$

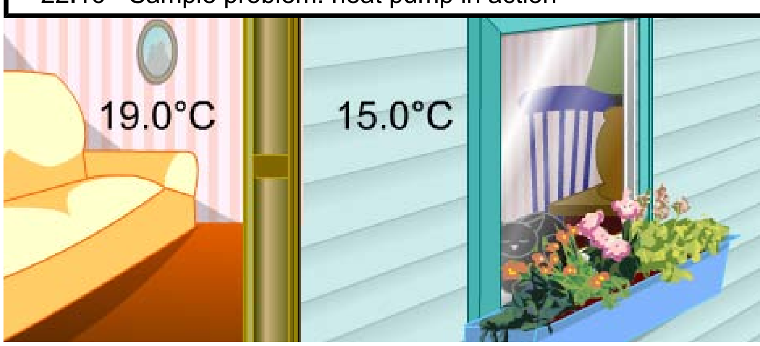
$T_h = 21 + 273.15\text{ K} = 294\text{ K}$

$T_c = 13 + 273.15\text{ K} = 286\text{ K}$

$$COP_{max} = \frac{294\text{ K}}{294\text{ K} - 286\text{ K}}$$

$COP_{max} = 37$

22.16 - Sample problem: heat pump in action



For these temperatures, what is the maximum *COP* for a heat pump? If the *COP* of a heat pump is 5.48, and 1250 J of work is done on it, how much heat energy could it maximally transfer inside from the outdoors?

If you heat a house using an electric furnace, the electric energy supplied to the furnace is converted into thermal energy and transferred to the air inside the house, increasing its temperature. If the furnace is supplied with 1250 joules of electrical energy, it can increase the energy of the air by at most 1250 J. It will be interesting to contrast this value with the amount of heat that a heat pump can cause to flow into the house using the same amount of energy.

Variables

temperature of cold reservoir	$T_c = 15.0°C$
temperature of hot reservoir	$T_h = 19.0°C$
work done on heat pump	$W = 1250$ J
maximum COP of pump	COP_{max}
actual COP of pump	$COP = 5.48$
heat flowing to interior	Q_h

What is the strategy?

1. Calculate the maximum coefficient of performance of the heat pump
2. Use the definition of coefficient of performance with the actual COP to calculate the maximum amount of heat transferred to the interior.

Physics principles and equations

This inequality specifies the maximum possible COP of a heat pump with given temperature reservoirs:

$$COP \le \frac{T_h}{T_h - T_c}$$

(Be sure to use Kelvin temperatures.)

The definition of COP

$$COP = Q_h/W$$

Step-by-step derivation

We first calculate the maximum possible coefficient of performance for the heat pump.

Step	Reason
1. $COP_{max} = \dfrac{T_h}{T_h - T_c}$	maximum COP
2. $COP_{max} = \dfrac{(19.0+273.15)\text{ K}}{(19.0+273.15)\text{ K} - (15.0+273.15)\text{ K}}$	substitute values
3. $COP_{max} = 73.0$	evaluate

The maximum possible COP is over ten times higher than the actual (and more realistic) COP of 5.48 as stated in the problem. We now calculate the heat transferred into the house, using this actual value for the coefficient of performance.

Step	Reason
4. $COP = Q_h/W$	definition of COP
5. $Q_h = (COP)W$	solve for Q_h
6. $Q_h = (5.48)(1250\text{ J})$	substitute values
7. $Q_h = 6850$ J	evaluate

The heat pump transfers significantly more heat than the furnace would. As you might expect, this does not violate the principle of conservation of energy. The device "pumps" thermal energy from outdoors to indoors, further cooling the outdoors in order to warm the indoors.

22.17 - Interactive summary problem: efficiency of an automobile engine

In the simulation to the right, you get to put an engine to work in a common situation. You are the driver of a car. You specify how much gasoline you want the engine to consume. When you press GO, the car will accelerate at full engine power until the gasoline is gone, at which point the simulation will stop and your calculations will be evaluated.

Your goal is to specify the amount of gasoline that will allow you to reach a speed of 90.0 km/h. Here, we will consider the engine as supplying an external force doing work on the car. The engine must do 9.35×10^5 J of work to accelerate the car to 90.0 km/h. This figure is a good approximation of the amount of work a real engine must do to both increase the car's kinetic energy to the desired amount and overcome forces such as air resistance, friction, and so forth.

To decide how much gasoline you need, you must first calculate the efficiency of the engine. You can do this by adding any amount of gasoline and pressing GO to see the resulting heat added and heat expelled while the engine runs. Use these values to calculate the engine's efficiency. Next, calculate how much heat needs to be added to the engine for it to perform the desired amount of work. The gasoline is the source of the added heat. To find out how much gasoline to add, you need to know that one liter of gasoline generates 1.30×10^{8} J of heat when combusted in the engine.

The correct amount of gasoline will be under a tenth of a liter. Calculate the amount of gasoline needed to reach 90.0 km/h to three significant digits (to a ten-thousandth of a liter), enter this value, and press GO to see the results. Press RESET to try again.

If you want, you could calculate the kinetic energy of the car at the final speed and determine what percent of the work goes to overcoming resistive forces versus increasing the KE of the car. (The mass of the car is 1280 kg.)

If you have trouble getting the right answer, see the section on engine efficiency.

22.18 - Gotchas

Using Celsius temperatures. All the equations in this chapter that contain a temperature factor require the temperature to be in Kelvin.

My friend has an engine that is 101% efficient. This is not possible: It contradicts the second law of thermodynamics. It also contradicts the first law. Making that claim probably breaks some FTC regulations too.

If I clean my room, it will become more ordered and I will be breaking the second law of thermodynamics. Not to dissuade you from cleanliness, but you are not breaking the law. Your room is not an isolated system; other parts of the system (like the molecules within your body) have become more disordered.

22.19 - Summary

The efficiency of a heat engine is the ratio of the net work it does divided by heat added during an engine cycle. There are several equivalent statements of the second law of thermodynamics. One states that no heat engine can transform 100% of the thermal energy supplied to it during a cycle into work.

Entropy is a property of a system. Entropy increases as heat is transferred into the system. The change in entropy can be calculated as the heat transferred during a reversible process (one in which the system can be returned to its initial state without additional energy) divided by the temperature. Another way to state the second law is that in any isolated system − such as the universe − entropy never decreases.

The maximum efficiency possible for any heat engine depends on the ratio of the temperatures of the cold and hot reservoirs. The greater the difference in temperatures, the lower the ratio and the more efficient the engine can be. A Carnot engine is a theoretical engine that achieves this maximum efficiency.

The internal combustion engine in most automobiles utilizes the Otto cycle. The efficiency of the internal combustion engine depends on the compression ratio, the ratio of the maximum to minimum gas volume in the engine.

Heat pumps can be used instead of furnaces to warm buildings. Unlike a heat engine, a heat pump uses work to transfer heat from the cool reservoir (outdoors) to the hot reservoir (the building interior). Just as we rated heat engines by their efficiency, for a heat pump the analogous quantity is its coefficient of performance. A heat pump's coefficient of performance increases when the difference in temperature between the hot and cold reservoirs decreases.

Statistical mechanics provides another view of entropy. Any system can be in a finite (although possibly very large) number of distinguishable configurations, and each configuration can be achieved by a number of equivalent microstates. The number of microstates making up a configuration is the configuration's multiplicity. Boltzmann's entropy equation states that the entropy of a system configuration is a constant times the natural logarithm of its multiplicity.

Equations

Efficiency

$e = W/Q_h$

$$e = 1 - \frac{Q_c}{Q_h}$$

$$e \le 1 - \frac{T_c}{T_h}$$

$$e_{ce} = 1 - \frac{T_c}{T_h} \text{ (Carnot engine)}$$

$$e_{se} = 1 - \frac{T_{min}}{T_{max}} \text{ (Stirling engine)}$$

$$e_{ice} = 1 - \frac{1}{(V_{max}/V_{min})^{\gamma-1}} \text{ (Otto cycle)}$$

Free expansion

$W = 0,\ Q = 0,\ \Delta E_{int} = 0$

Entropy

$$\Delta S = \frac{Q_{rev}}{T}$$

In an isolated system, $\Delta S \ge 0$

Coefficient of performance

$$COP = \frac{Q_h}{W} = \frac{1}{e}$$

$$COP \le \frac{T_h}{T_h - T_c}$$

Chapter 22 Problems

Conceptual Problems

C.1 An amount of gas in a well-insulated container is compressed reversibly and adiabatically to half its initial volume. Does the entropy of the gas increase, decrease, or stay the same?

i. Increases
ii. Decreases
iii. Stays the same

C.2 A cup of hot coffee is placed inside a large, well-insulated container. Over a five minute period, the entropy of the cup decreases by 100 J/K. What can you say about the entropy of the air surrounding the cup?

i. It does not change
ii. It increases less than 100 J/K
iii. It increases by exactly 100 J/K
iv. It increases by more than 100 J/K
v. It decreases

C.3 A refrigerator operates much like a heat pump, doing work to transfer heat from its interior to its exterior. If you leave a refrigerator running with its door open in the middle of a well-insulated room, does the room temperature go up, go down, or remain unchanged? Explain.

The room temperature
i. goes up
ii. goes down
iii. is unchanged

C.4 In a reversible thermodynamic process, the entropy of a system increases but the temperature is unchanged. What type of process is this?

i. Constant-volume
ii. Isobaric
iii. Isothermal
iv. Adiabatic

C.5 In a reversible thermodynamic process, a system's entropy is unchanged, while its temperature increases. What type of process is this?

i. Constant-volume
ii. Isobaric
iii. Isothermal
iv. Adiabatic

C.6 Which of the following statements must be true of system that undergoes a reversible adiabatic process?

i. There is no work done by the system.
ii. The entropy of the system is unchanged.
iii. The temperature of the system is unchanged.
iv. The internal energy of the system is unchanged.

Section Problems

Section 1 - Efficiency

1.1 An engine running at 38.3% efficiency receives 287 J of heat from the hot reservoir during each cycle. How much heat is transferred to the cold reservoir during each cycle?

__________ J

1.2 An engine does 189 J of net work during a cycle, using 329 J of heat in from the hot reservoir in the process. What is the engine's percent efficiency?

__________ %

1.3 A two-stroke engine runs at 2300 rpm, which means it performs 2300 engine cycles each minute. If the engine does 9.2×10^5 J of net work each minute, and has an efficiency of 0.43, how much heat is transferred into the engine during a cycle?

__________ J

1.4 During one engine cycle, a lawnmower engine does 55.2 J of net work. In the process, it expels 153 J of energy as heat. What is the percent efficiency of the engine?

__________ %

Section 5 - Entropy

5.1 In a reversible process called isothermal expansion, 457 J of heat are transferred into a quantity of gas in an engine from the hot reservoir, while 352 J of heat are transferred out to the cold reservoir. The temperature of the gas is 473 K at the beginning and end of the process. What is the change in entropy of the gas?

__________ J/K

Section 8 - Second law of thermodynamics: entropy

8.1 Two large blocks are momentarily placed in contact. Block A has temperature 567 K and transfers 849 J of heat energy to block B, which has temperature 288 K. The temperatures of the blocks are unchanged. What is the change in entropy of the system consisting of the two blocks?

__________ J/K

8.2 An ice cube of mass 0.018 kg, at the temperature of the freezing point of water, is surrounded by water at a temperature just above the freezing point. What is the change in entropy of the ice cube as it melts? (The latent heat of fusion of water at the freezing point is 3.34×10^5 J/kg.)

__________ J/K

8.3 An ideal gas expands isothermally at 64.0°C in a reversible process, and its entropy increases 2.00 J/K. What amount of heat is transferred to the gas in the process?

__________ J

Section 10 - Maximum engine efficiency and reservoir temperatures

10.1 An engine operates with a cold reservoir at 148°C. If the design goal is to achieve an operating efficiency of 42.4%, what is the minimum temperature required for the hot reservoir, in degrees Celsius?

__________ ° C

10.2 An engine, built by an advanced alien civilization, operates at the maximum possible efficiency for the temperatures of its hot and cold reservoirs. The engine's efficiency is 37.2% and the difference between the temperatures of the two reservoirs is 93.0 K. (a) What is the temperature of the hot reservoir? (b) What is the temperature of the cold reservoir?

(a) hot reservoir __________ K
(b) cold reservoir __________ K

Section 12 - Carnot cycle and efficiency

12.1 A Carnot engine operates with a cold reservoir temperature of 194 K and has an efficiency of 74.3%. (a) What is the temperature of the hot reservoir? (b) If the engine does 8.99×10^4 J of net work during each cycle, how much heat does it absorb from the hot reservoir during a cycle?

(a) __________ K
(b) __________ J

12.2 A Carnot engine absorbs 4.65×10^3 J from the hot reservoir and expels 3.12×10^3 J to the cold reservoir during each engine cycle. If the temperature of the cold reservoir is 489 K, what is the temperature of the hot reservoir?

__________ K

12.3 A Carnot engine operates with reservoirs at 135°C and 378°C, and 2.38e+4 J of heat are transferred into the engine from the hot reservoir during each engine cycle. How much net work does the engine do per cycle?

__________ J

12.4 A Carnot engine operating with reservoirs at 459 K and 726 K has a power output of 7.35×10^4 W. (a) How much heat energy does it absorb from the hot reservoir each hour? (b) How much heat energy does it expel to the cold reservoir each hour?

(a) __________ J
(b) __________ J

12.5 An engine removes 734 J from the hot reservoir, which is at 487 K, and expels 523 J of heat to the cold reservoir, which is at 228 K. (a) What is the engine's percent efficiency? (b) What would the percent efficiency of a Carnot engine operating with these reservoirs be?

(a) __________ %
(b) __________ %

12.6 An engine with reservoir temperatures of 387 K and 725 K operates at half the efficiency of a Carnot engine. If the hot reservoir supplies the engine 4.72e+3 J of heat energy during each cycle, how much net work does the engine do in a cycle?

__________ J

Section 13 - Otto cycle: internal combustion engine

13.1 An internal combustion engine has a compression ratio of 4.50 and the ratio γ of the molar specific heats of the air-fuel mixture is 1.43. What is the maximum efficiency of the engine, as a percent, if it operates in an ideal Otto cycle?

(a) __________ %

13.2 An engineer is designing an internal combustion engine. The air-fuel mixture used in the engine has a γ ratio of 1.37. Each cylinder of the engine has a maximum volume of 4.12×10^{-4} m^3. If the desired efficiency of the engine, operating in an ideal Otto cycle, is 59.3%, what should the compressed volume of the air-fuel mixture in the cylinder be?

__________ m^3

13.3 An internal combustion engine has a maximum ideal efficiency of 33.6%, based on its compression ratio and the characteristics of the air-fuel mixture. (a) If the compression ratio is 6.45, what is the value of γ for the gas? (b) To increase the efficiency to 46%, what should the value of γ be?

(a) __________
(b) __________

13.4 In an internal combustion engine, the air-fuel mixture at the end of the ignition phase has volume 7.34×10^{-4} m^3 and pressure 3.12×10^{6} Pa. The air-fuel mixture expands adiabatically, pushing the piston out, to a volume of 2.82×10^{-3} m^3. Assume the air-fuel mixture behaves like an ideal gas. If γ = 1.42 for the air-fuel mixture, what is the final pressure after the expansion?

__________ Pa

Section 15 - Heat pumps

15.1 A Carnot engine operated "in reverse" as a heat pump achieves the maximum coefficient of performance. A Carnot heat pump operates to warm a house, where the outside temperature is 12°C and the inside temperature is 31°C. (a) What is its *COP*? (b) How much work must be done on the heat pump for it to deliver 1400 J of heat energy to the interior of the house?

(a) __________
(b) __________ J

15.2 A heat pump on a house operates with a *COP* that is 35% of the theoretical maximum, between exterior and interior temperatures of −12°C and 34°C. How much work must be done on the heat engine to supply 2200 J of heat to the interior of the house?

__________ J

15.3 A refrigerator operates much like a heat pump, but the desired goal of a refrigerator is to transfer heat from the cold reservoir as efficiently as possible. For this reason, the coefficient of performance of a refrigerator is measured as Q_c/W. A typical commercial refrigerator has a coefficient of performance of about 5.5. Assume such a refrigerator is used by a market to make ice cubes from 2.3 kg of liquid water, starting at 0°C. (a) How much work is done on the refrigerator to freeze this water into ice? (The latent heat of fusion for the water is 3.34×10^{5} J/kg.) (b) Suppose water costs 0.56 cents per kg, and energy costs 10 cents per kWh. If the market charges 99 cents for 2.3 kg of ice, what is its maximum profit on the ice? Express your answer to the nearest cent.

(a) __________ J
(b) __________ cents

Section 17 - Interactive summary problem: efficiency of an automobile engine

17.1 Use the information given in the interactive problem in this section to calculate the correct amount of gasoline. Test your answer using the simulation.

__________ L

Additional Problems

A.1 An ideal gas expands in a reversible process at temperature 287 K, from 456 m^3 to 497 m^3. The number of moles of gas is 2.12. (a) Calculate the entropy change of the gas if the expansion is isothermal. (b) What is the entropy change if the expansion is adiabatic instead of isothermal?

(a) __________ J/K
(b) __________ J/K

A.2 The hot reservoir of an engine operates at 472 K and the cold reservoir at 214 K. The engine recieves 433 J of heat from the hot reservoir during a cycle. The engine's efficiency is 67.2% of the maximum possible (for a Carnot engine). (a) What is the engine's efficiency, as a percent? (b) How much net work does the engine do during a cycle? (c) How much heat is absorbed by the cold reservoir each cycle?

(a) __________ %
(b) __________ J
(c) __________ J

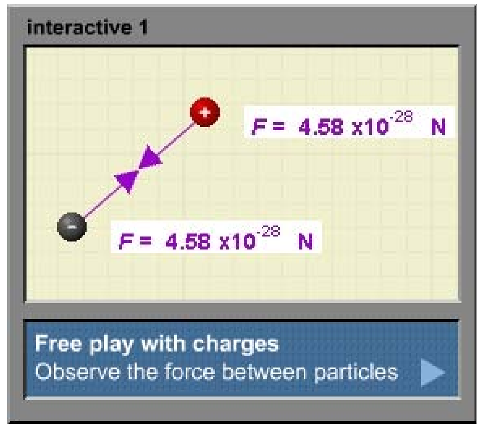
A.3 In a reversible thermodynamic process, 780 J of heat is transferred from the hot reservoir, with temperature 658 K, to a cold reservoir, with temperature 362 K. (a) What is the resulting change in the entropy of the universe? (b) What is the net work done during this time?

(a) __________ J/K
(b) __________ J

chapter 23 Electric Charge and Coulomb's Law

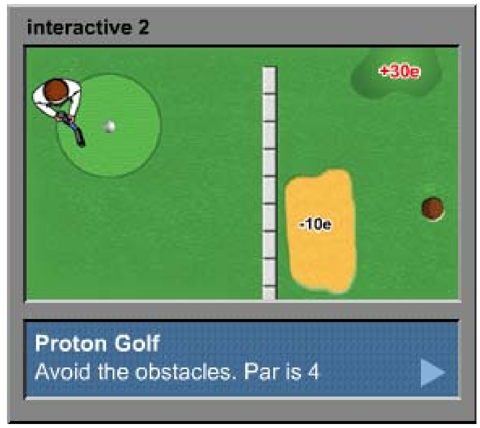
23.0 - Introduction

In this chapter, we begin the study of electricity and magnetism by discussing electric charge and the electrostatic force. Although you cannot see the individual charged particles, such as electrons and protons, that cause this force, you certainly see its effects. Phenomena ranging from the annoying static cling in freshly laundered clothing to the operation of laser printers are based on the electrostatic force. In the sections that follow, we will cover the fundamentals of electric charge: what it means to say that an object is charged and the nature of the forces created by charged objects.



We start with two simulations, shown to the right. The first allows you to experiment with positively and negatively charged particles and see the forces they exert on one another. The positively charged particles in this simulation have the same charge as protons, and the negatively charged particles have the same charge as electrons.

After you launch this simulation, drag particles from the control panel onto the screen above it. Once there, they will exert forces on each other. The amount and direction of each force will be shown on the screen. You can drag particles closer together or farther apart to see how the force they exert on one another relates to the distance between them (the heavier grid lines are exactly one meter apart). If you press GO, the particles, which all have equal mass, will be free to accelerate in response to the forces they exert on each other. *Electrostatics* is the study of electric charges at rest, so the simulation is also providing you with an extremely informal introduction to the topic of *electrodynamics*, the study of charges in motion.



As you use the simulation, take note of the direction of the forces between, say, two negative or positive particles or between a positive and a negative particle. You can also place two particles with the same charge next to each other, and see how the force on a third particle changes. How the electric force changes with both the distance between the charged particles − frequently just called "charges" − and the amount of charge is a fundamental focus of this chapter.

In the second simulation, you can play "proton golf". The ball is positively charged, and you add protons to the putter to make it positively charged as well. The protons in the putter exert a force, called the electrostatic force, on the ball even when the two are not touching. You can control both the location of the putter, by dragging it, and how many protons it contains, by clicking the up- and down-arrows in the console. The moment you load protons into the putter they exert a force on the ball, but the ball is locked in place until you press PUTT. The grass of the green supplies a frictional force that will cause the ball to stop rolling.

Your mission, as always in golf, is to sink the ball in the hole − in four or less shots, if you can! The important thing (in addition to having fun) is to observe how the electrostatic force relates to the amount of charge and to distance. Be warned, though: Obstacles do exist! A clump of protons acts as a hill that causes the ball to roll away from it, while a clump of electrons is a sand trap that will attract the ball. Fore!

23.1 - Electric charge

Electric charge: A property of the particles that make up matter. It causes attraction and repulsion.

This woman's hair is electrically charged. As you will see, the strands of her hair repel each other because each one of them carries a negative charge.

Electric charge is a property of matter that can cause attraction and repulsion. In this section, we focus on electrons and protons, and the role they play in causing an object to have an electric charge.

An electron is defined as having a **negative** charge and a proton is defined as having a **positive** charge. Charge is a scalar, not a vector. A negative charge is not less than zero, just the opposite of positive. In this book we will represent negative charges as black and positive charges as red.

The amount of charge of an electron or proton is written as e and is called the *elementary charge*. An electron has a negative charge of $-e$ and a proton has a positive charge of $+e$. This amount of charge is the smallest amount that has been isolated. (Subatomic particles called *quarks* have charges of $+2e/3$ or $-e/3$ but they have not been isolated.)

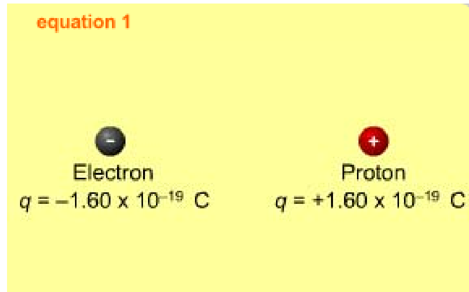
The SI unit for charge is the *coulomb.* An electron or a proton has a charge of magnitude e = 1.602 18×10^{-19} coulombs. This means approximately 6,250,000,000,000,000,000 electrons or protons are required for a coulomb of charge to be present. This is a vast number! However, numbers like this are often present in nature: A bolt of lightning typically contains about 25 C of charge. To provide you with another idea of the magnitude of a coulomb, approximately 0.8 C of charge flows through a 100 watt light bulb every second.

Some scientists, chemists in particular, use another unit, the *esu* or *electrostatic unit*. One esu equals $3.335\,64\times10^{-10}$ C.

A small amount of matter contains a large number of electrons and protons. For instance, a one-kilogram sample of copper contains about 2.75×10^{26} protons. When an object has the same number of electrons and protons, it has no net charge and is said to be *electrically neutral*.

The addition or removal of electrons from an object causes it to become charged. A negatively charged object has more electrons than protons and a positively charged object has more protons than electrons. If the kilogram of copper has a charge of +0.1 C, which is a relatively large amount of charge, this means that about 0.000 002 % of its electrons have been removed.

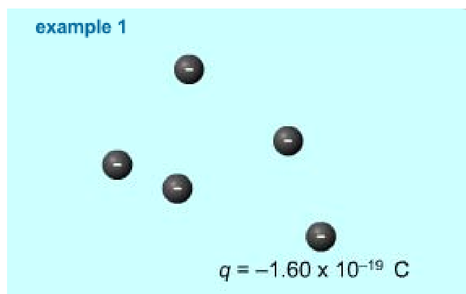
Electric charge
Property of particles that make up matter
· Electrons negative, protons positive
· They have opposite amounts of charge



Electric charge

Electron: $q = -1.60\times10^{-19}$ C
Proton: $q = +1.60\times10^{-19}$ C

q is symbol for charge
Units: coulombs (C)



How much charge do these five electrons have?

q_{total} = (5) (charge of 1 electron)
electron charge = -1.60×10^{-19} C
$q_{total} = (5)\,(-1.60\times10^{-19})$
q_{total} = **negative** 8.00×10^{-19} C

23.2 - Creating charged objects

How does an object become electrically charged? The answer is that the addition or removal of electrons creates negatively and positively charged objects. Except under extreme conditions, protons stay in place and electrons move.

A piece of silk and a glass rod can be used to demonstrate one manner in which objects can become charged. We will assume these two objects start out electrically neutral. In other words, the silk has equal numbers of protons and electrons, as does the glass.

You can transfer electrons from the glass to the silk by rubbing the two materials together. This close contact results in a net flow of electrons from the glass to the silk and causes the silk to become negatively charged. It now contains more electrons than protons. In turn, the glass becomes positively charged, since it now has fewer electrons than protons.

How does an object become electrically charged? The answer is that the addition or removal of electrons creates negatively and positively charged objects. Except under extreme conditions, protons stay in place and electrons move.

A charged comb induces electric charges in the paper dots which cause them to stick together. This phenomenon is called static cling.

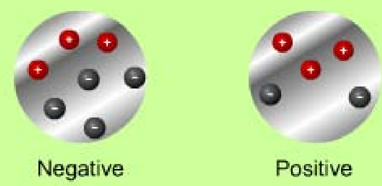
A piece of silk and a glass rod can be used to demonstrate one manner in which objects can become charged. We will assume these two objects start out electrically neutral. In other words, the silk has equal numbers of protons and electrons, as does the glass.

You can transfer electrons from the glass to the silk by rubbing the two materials together. This close contact results in a net flow of electrons from the glass to the silk and causes the silk to become negatively charged. It now contains more electrons than protons. In turn, the glass becomes positively charged, since it now has fewer electrons than protons.

You may wonder why rubbing silk and glass together causes them to become charged. The electrons move because the silk molecules have a greater affinity for electrons than do the glass molecules. Rubbing the two materials together facilitates the transfer of electrons by providing a greater level of contact between their molecules.

The charging process can be reversed. When free to move, electrons will flow from a negatively charged object to a positively charged one, reducing or ending a charge imbalance. Lightning provides a dramatic example of such movement, a grand display of excess electrons moving to a region that is less negatively charged. With lightning, the electrons may be moving to a positively charged region of a cloud, or to an electrically neutral region such as the surface of the Earth. Charges take advantage of any opportunity to reduce an imbalance.

concept 1



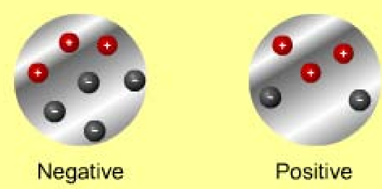
Creating charged objects

Neutral objects become charged by movement of electrons

Excess electrons: negatively charged

Excess protons: positively charged

equation 1



Creating charged objects

$$q = \pm Ne$$

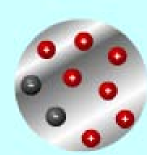
q = charge

N = number of excess charges

· protons positive, electrons negative

e = elementary charge

example 1



What is the net charge on the

disk?

N = 7 protons – 2 electrons

N = 5 protons

$q = +Ne$

$q = (5)(1.60\times10^{-19})$

$q = 8.00\times10^{-19}$ C

23.3 - Conservation of charge

Conservation of charge: The net charge of an isolated system of objects remains constant.

Although lightning transfers a very large amount of charge from a cloud to the ground, the total charge remains constant.

Electric charge is conserved. The net charge of an isolated system may be positive, negative or neutral. Charge can move between objects in the system, but the net charge of the system remains unchanged.

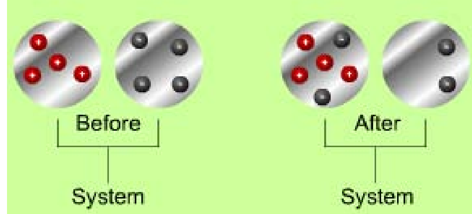
To illustrate this principle, we again use the example of a silk cloth and a glass rod to demonstrate how two different objects can become charged while, at the same time, overall charge is conserved. Let's assume that the cloth and the rod are both neutral to begin with. They each become charged when rubbed together, but their combined charge is unchanged: It remains zero, or neutral. It is true that electrons have moved between the rod and the cloth, but to the extent that one object is negative, the other is positive. The cloth and the rod constitute an isolated system because all the charge moves solely between them and no charge leaves them. If charge flowed to a person holding these objects, that person would become part of the system as well, and charge would still be conserved.

Early on, scientists such as Benjamin Franklin (yes, *that* Ben Franklin) suggested this conservation principle based on experimental data and intuition, but he and his colleagues were unable to show why it was so. The discovery of electrons showed why the conjecture was true: It was the movement of electrons that created the charged objects that Franklin observed. Under ordinary circumstances, these particles are neither created nor destroyed, and Franklin observed the results of electrons flowing from one object to another.

An object is charged when it has an imbalance of electrons and protons. Charge is said to be *quantized*: It is always observed as an integer multiple of e, the magnitude of the charge of an electron or a proton.

In extreme circumstances, charged particles can be destroyed. For example, when a positron (an exotic particle that is the mirror image of an electron, identical in mass but opposite in charge) collides with an electron, the two will annihilate each other and produce gamma rays, a kind of high-energy radiation. Does this scenario violate the conservation of charge? No, because gamma rays have no net charge. Before the collision, the system of one electron and one positron has no net charge. After the collision, the system consists of neutral gamma radiation, so charge is conserved.

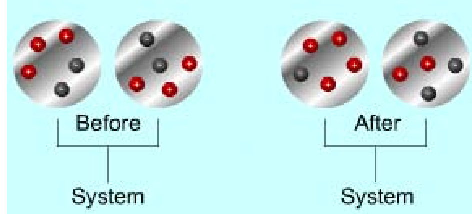
concept 1



Conservation of charge

Charges can move but system's net charge is constant

example 1



Is charge conserved in this system?

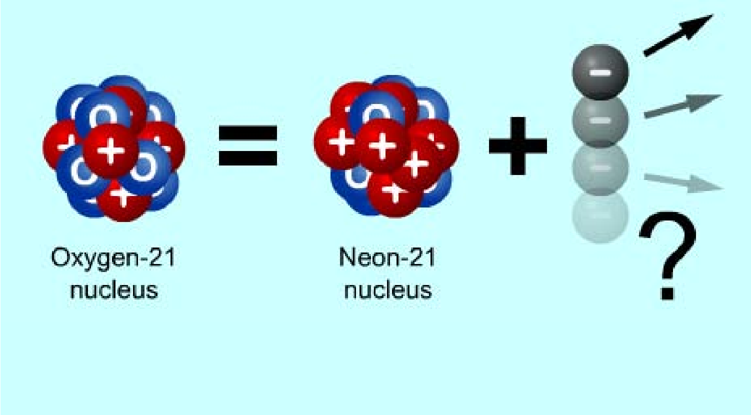
Before: 2 excess protons

After: 2 excess protons

Net charge before = net charge after

Charge conserved!

23.4 - Sample problem: charge conservation in nuclear decay



The nucleus of an atom of oxygen contains 8 protons, while the nucleus of an atom of neon contains 10 protons. Via a process called beta decay, a neutron can decompose into one proton plus one electron, and oxygen-21 transmutes into neon-21.

Applying conservation of charge, how many electrons must be emitted by the oxygen nucleus as the atom transmutes into neon-21?

Variables

charge of oxygen-21 nucleus	$+8e$
charge of neon-21 nucleus	$+10e$
charge of one electron	$-e$

What is the strategy?

1. Calculate the change in the positive charge of the nucleus when it transmutes into neon.
2. The number of electrons emitted must balance this change in charge.

Physics principles and equations

The conservation of charge holds for all isolated systems. A process like beta decay will not change the total charge of an isolated system.

Solution

To transmute an atom of oxygen into neon, two additional protons are required in the nucleus. The change in charge is $+10e - (+8e)$, which is $+2e$. Since charge is conserved, the total charge does not change, and the process of beta decay must also result in the creation of a charge of $-2e$, which is the charge of two electrons. The nucleus emits the two electrons, and the total charge of the system containing the atom and its emissions does not change.

Oxygen-21 is what is called an unstable isotope (form) of oxygen. The beta decay process described in this problem occurs in two steps: first oxygen-21 transmutes to the unstable fluorine-21 and then fluorine-21 transmutes to the stable substance neon-21.

23.5 - Conductors, insulators, and grounds

Conductor: An object or material in which charge can flow relatively freely.

Insulator: An object or material in which charge does not flow freely.

The lightning rod mounted atop this cupola intercepts lightning and protects a building during electrical storms.

Ground: Charge flows from a charged object to a ground, leaving the object neutral.

You can easily find conductors and insulators (also called *nonconductors*) in your home or classroom. If you examine an electrical cord, you will find that it consists of a conducting core of copper wire surrounded by an insulator such as vinyl plastic.

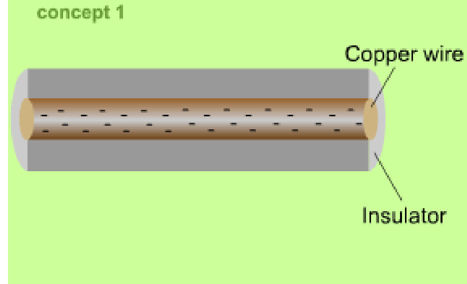
Charge can be moved relatively easily through a conductor such as copper using a device like a battery. A battery will cause electrons to flow through a copper wire like the one shown in Concept 1. In contrast, it is difficult to cause electrons to flow in insulators like rubber or many plastics. This difference explains the design of electrical cords: Often, they are made of copper wire wrapped with a flexible vinyl insulator so that electrons flowing through the wire remain within the cord.

Insulators do not allow charges to flow when they are subjected to only moderate amounts of force. When great amounts of force are applied, charge can flow through an insulator. There are also materials called semiconductors that enable charge to flow in some circumstances, but not others. Given their role in devices like transistors, they are an important topic, but lie outside the scope of this section.

A ground is a neutral object that can accept or supply an essentially unlimited number of charges. The Earth functions as an electric ground. If you touch a conducting, charged object to the ground, the object will also become electrically neutral − in other words, grounded. Excess electrons will flow out of a negatively charged object to the ground, and electrons will flow into a positively charged object from the ground. Charges move to a ground because charges of the same sign move as far away from each other as possible due to their mutual repulsion. The ground distributes the excess charge far enough away that it ceases to affect the object.

Protecting houses from lightning presents engineers with the need to use conductors and grounds. A building is not usually a conductor, but lightning can transform a house into a reluctant conductor, with disastrous, highly flammable results. A lightning rod is a conductor that protects houses and other structures by providing an easier, alternate route to the ground. A conducting wire connects the rod to the Earth. The shape of the rod also increases the likelihood that it will be the preferred target for lightning.

concept 1



Conductors and insulators
Conductor: charge moves freely
Insulator: charge does not readily flow

concept 2

Ground
Makes conductors electrically neutral

23.6 - Interactive problem: charged rods

These interactive simulations are versions of a classic game. In the original version of this game, you are given glasses, some filled with water and others empty. You are shown or told a final configuration of glasses and water. Your challenge is to start with the initial configuration, and by pouring water from one glass to another in a sequence of steps, end with the specified final configuration. For instance, a simple challenge would be to start with an empty glass, a half-full glass, and a quarter-full glass, and end up with a three-quarters-full glass. By pouring the half-full glass into the quarter-full glass, you achieve that goal.

In the simulations to the right, the same overall idea applies to electric charge. You are supplied with a configuration of charges on rods. Some of the rods have no charge, some have positive charge, and some are negatively charged. All the rods are the same size and are made of identical conducting material. In this game, charge flows between rods instead of water flowing between glasses. Charge flows until equilibrium is reached. For example, if you touch a rod with +4.000 microcoulombs of charge to a rod with no charge, both rods end up with +2.000 microcoulombs of charge.

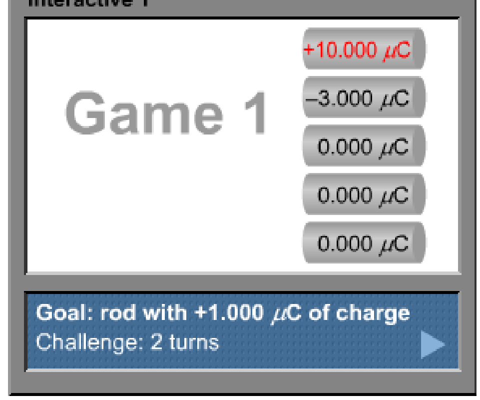
To play the game, click on any rod and drag it to another rod. When you release the mouse button, charge will transfer between the rods.

In the topmost game to the right, you are given a rod with a charge of positive 10.000 μC, a rod with a charge of −3.000 μC, and several neutral rods. Your goal is to produce a rod with a charge of +1.000 μC. This can be done in two moves. The second game requires a greater number of moves and more planning. You can see the initial configuration to the right. The challenge again is to create a rod with +1.000 μC of charge. However skilled you are at these two games, the main point is to observe how charge is conserved.

Keep in mind that you do not have to be good at the games to practice the physics you are learning in this chapter. Give it a try! Whether you get the minimum number of moves or not, the simulations offer a chance to employ the principle of conservation of charge.

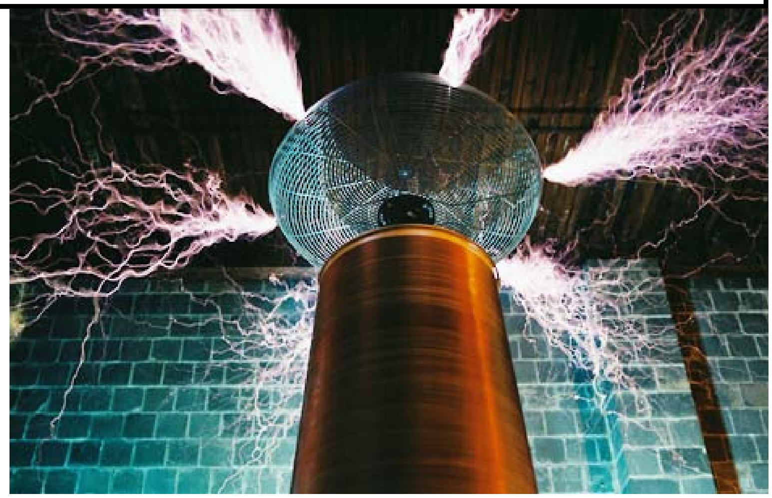
If you have any questions about the conservation of charge or about grounds, review the preceding sections on these topics.

interactive 1



interactive 2

Electrostatic force: Attraction or repulsion due to electric charge.



This Van de Graaff generator builds up an enormous electrostatic charge that escapes into the surrounding air.

Electrostatics is the study of electric forces between charges at rest. If you have ever visited a science museum, you may have seen people press their hands against an electrically charged device surmounted by a shiny metallic sphere, and then watched in amazement as their hair stands on end. This device, called a *Van de Graaff generator*, amusingly illustrates how electric charge creates a repulsive force. In the photograph above you see the spectacular display that can be created by such forces in a large Van de Graaff generator, as its huge electrostatic accumulation discharges through the atmosphere. The Boston Museum of Science states that it is home to the largest Van de Graaff generator in the world.

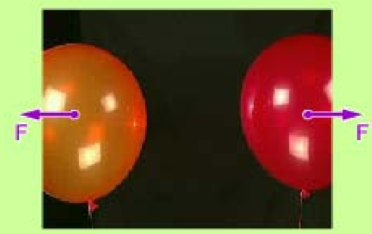
Clothes dryers provide a more mundane example of electrostatic forces at work. When your socks stick to your pants and then crackle as you pull them apart, you are witnessing the static cling caused by electrostatic forces. In this case, electrostatic force is causing oppositely charged pieces of clothing to attract each other. As the clothing is pulled apart, electric charges arc between the clothing items in an attempt to reach a more balanced state. (Imagine: When you fold your laundry, you can both please your parents and review your physics studies. What a deal!) The electric charge responsible for that annoying cling in a sock is typically in the range of a few microcoulombs.

When objects have opposite charges, like laundry items or a glass rod and a silk cloth, they attract. When objects like the two balloons you see to the right have the same charge, either positive or negative, they repel each other. The old cliché − opposites attract and likes repel − proves true in physics. When dealing with issues of attraction and repulsion, it really is important to know your sign.

Two charged objects exert equal but opposite forces on each other. In other words, if they attract, they pull toward each other with the same force. If they repel, they push against each other with equal force.

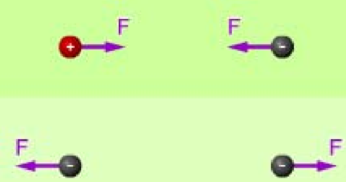
The forces act along a straight line between the centers of the two charges. For instance, if they attract, each force points directly toward the other charge, as illustrated in Concept 2. If they repel, each force points directly away from the other charge.

concept 1



Charged balloons repel each other
Charged objects can attract or repel

concept 2



Electrostatic force
Opposites attract
Likes repel

example 1

Will the balloons attract or repel?
Same charge
Balloons repel

Inducing an electric charge: Creating a charged object or region of an object without direct contact.

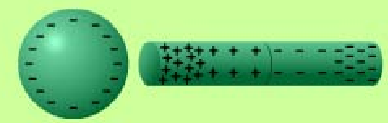
Objects can become electrically charged when they are put into contact with each other, for example, by rubbing glass and silk together, or by touching a charged rod to a neutral one. In this process electrons flow from one object to the other.

Objects can also become charged without touching. Like gravity, electrostatic forces act at a distance, so charges cause other charges to move without direct contact. When a charged object, like the nonconducting sphere shown in Concept 1, is placed near a neutral object in which electrons are free to move, such as the joined pair of conducting metal rods to its right, the charged object causes electrons to move in the neutral object. Charges in the rods, initially evenly distributed throughout the pair, end up in the asymmetrical configuration you see in Concept 1. The rod pair as a whole is still electrically neutral.

To explain how charged objects can be created without direct contact, we use the sphere and the pair of neutral rods just discussed. First, the negatively charged sphere approaches the rods. As the sphere repels electrons in the rods, the closer end of the rod combination becomes positively charged. As a result of the movement of electrons, the far end of the rod combination becomes negatively charged. Two regions of charge have been induced without contact by a charged object.

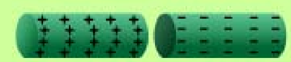
Next, the rods are separated. The closer one will remain positive and the farther one will remain negative, even after the charged sphere has moved away. This example shows how two objects can become charged without coming into direct contact with a third charged object.

concept 1



Inducing an electrical charge
Creating charged objects without direct contact
· Charge on ball moves electrons in rod pair to create positive, negative regions

concept 2



Inducing an electrical charge
· Separating rods completes induction

23.9 - Coulomb's law: calculating electrostatic forces

Coulomb's law: The electrostatic force a charged particle exerts on another is proportional to the product of the charges and inversely proportional to the square of the distance between them.

Named for Charles Augustin de Coulomb, the eighteenth century French physicist who formulated it, this law quantifies the amount of force between charged particles. His law is shown in Equation 1 to the right.

The force is measured in newtons. The constant k in the equation has been experimentally determined. It equals 8.987 55×10^9 N·m^2/C^2.

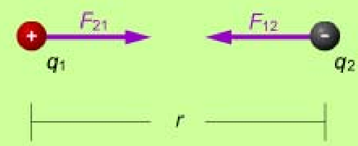
The charges are shown with absolute value signs around them, so that two positive values are multiplied together to calculate the **amount** of the force. To determine the direction, the rule "opposites attract, likes repel" is used. Two opposite charges attract, so both forces pull the charges together. Two like charges repel, so both forces push the charges away from each other. Recalling Newton's third law helps to insure the correct result: The forces are always equal but opposite to each other.

Coulomb's law means that larger quantities of charge create more force and that the force weakens with the square of the distance.

Electrostatic forces can be added; they obey the principle of superposition. For example, if there are three charges, the force exerted by two of the charges on the third equals the vector sum of the forces exerted by each charge. This is similar to other forces you have studied; if two people are pushing a car, the net force equals the vector sum of the individual forces exerted by each person.

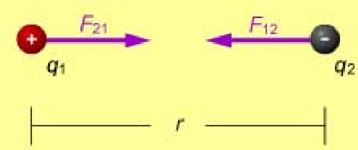
In Coulomb's law, r is the distance between two *point charges*, two infinitesimal sites of charge. If charges are symmetrically distributed on each of two spheres, a principle called the shell theorem can be used to show that all the charge on each sphere acts as

concept 1



Coulomb's law
Electrostatic force
· Proportional to product of charges
· Inversely proportional to distance squared

equation 1



Coulomb's law

though it were located at the sphere's center. In this case, the distance r is the distance between the centers of the spheres.

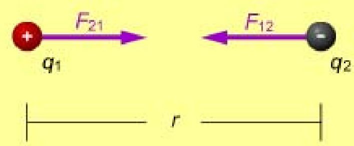
If you have studied gravity, you may notice that Coulomb's law is similar to the equation for calculating gravitational force. Both are *inverse square laws*: The forces are inversely proportional to the distance squared. With Coulomb's law, the force is proportional to the product of the charges; with gravity, the force is proportional to the product of the masses. Both are field forces, acting at a distance. Similarities like these cause physicists to search for one unified explanation of gravitational and electrostatic forces. Do remember, however, there is a crucial difference between the two forces: Masses always attract, while electric charges can both attract and repel.

Sometimes Coulomb's law is expressed in another fashion, using the *permittivity constant* ε_0. This traditional way of expressing the law can be particularly helpful in your later studies. The equation expressed with the permittivity constant is also shown to the right, as Equation 2. The permittivity constant is related to Coulomb's constant by the equation $\varepsilon_0 = 1/4\pi k$, and it equals 8.854 19×10^{-12} C^2/N·m^2.

$$F = k\frac{|q_1||q_2|}{r^2}$$

F = force
k = Coulomb's constant
q = charge
r = distance between charges
Constant $k = 8.99\times10^9$ N·m^2/C^2

equation 2

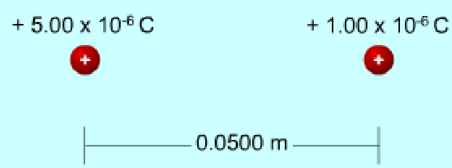


Coulomb's law, permittivity constant

$$F = \frac{1}{4\pi\varepsilon_0}\frac{|q_1||q_2|}{r^2}$$

ε_0 = permittivity constant
Constant $\varepsilon_0 = 8.85\times10^{-12}$ C^2/N·m^2

example 1



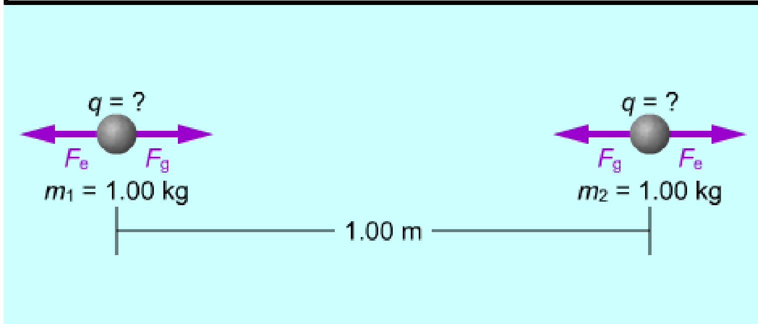
What is the magnitude and direction of the force the right charge exerts on the left charge?

$$F = k\frac{|q_1||q_2|}{r^2}$$

$$F = k\frac{(5.00\times10^{-6}\,\mathrm{C})(1.00\times10^{-6}\,\mathrm{C})}{(0.0500\ \mathrm{m})^2}$$

F = 18.0 N (to the left)

23.10 - Sample problem: electric vs. gravitational force



How many excess electrons must be added to each neutral lead sphere to balance the force of gravity between them?

The two balls above are floating in deep space, with the only significant gravitational forces acting upon them being the ones they exert upon one another. If they were electrically neutral, they would drift slowly together due to these forces and, after about an hour, come to rest against each other.

You are asked to determine how many electrons should be added to each sphere so that the electrostatic force exactly counteracts the gravitational force. You add the same number of electrons to each sphere, and disregard the change in mass of the spheres due to the added electrons. It is negligible.

Variables

	sphere 1	sphere 2
mass of sphere	$m_1 = 1.00$ kg	$m_2 = 1.00$ kg
charge of sphere	q_1	q_2
gravitational constant	$G = 6.67\times10^{-11}$ N·m²/kg²	
distance between spheres	$r = 1.00$ m	
gravitational force on sphere	F_g	
electric force on sphere	F_e	
Coulomb's constant	$k = 8.99\times10^{9}$ N·m²/C²	
number of excess electrons	N	
elementary charge	$e = 1.60\times10^{-19}$ C	

What is the strategy?

1. Calculate the gravitational attraction between the lead spheres.
2. Set the gravitational force equal to the repulsive electrostatic force between them and solve for the charge. The amount of charge on each sphere is the same.
3. Convert the charge from coulombs to the equivalent number of excess electrons.

Physics principles and equations

We will use Newton's law of gravitation

$$F_g = G\frac{m_1 m_2}{r^2}$$

together with Coulomb's law

$$F_e = k\frac{|q_1||q_2|}{r^2}$$

The charge due to N excess electrons is

$$q = -Ne$$

Step-by-step solution

In the first steps, we calculate the charge q needed on the spheres to balance their gravitational attraction.

Step	Reason
1. $F_g = G\dfrac{m_1 m_2}{r^2}$	Newton's law of gravitation
2. $F_g = G\dfrac{(1.00\text{ kg})(1.00\text{ kg})}{(1.00\text{ m})^2}$ $F_g = 6.67\times 10^{-11}\text{ N}$	evaluate
3. $F_g = F_e = k\dfrac{\lvert q\rvert^2}{r^2}$	set forces equal and use Coulomb's law
4. $q = \pm\sqrt{\dfrac{F_g r^2}{k}}$	solve for q
5. $q = -\sqrt{\dfrac{(6.67\times 10^{-11}\text{ N})(1\text{ m})^2}{8.99\times 10^{9}\text{ N}\cdot\text{m}^2/\text{C}^2}}$ $q = -8.61\times 10^{-11}\text{ C}$	note electron charge negative and evaluate

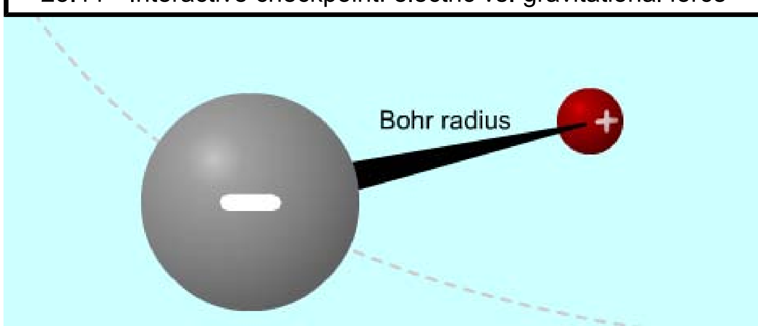
In the following steps we find the number of excess electrons.

Step	Reason
6. $q = -Ne$	equation for charge
7. $N = -\dfrac{q}{e}$	solve for N
8. $N = -\dfrac{-8.61\times 10^{-11}\text{ C}}{1.60\times 10^{-19}\text{ C}}$ $N = 5.38\times 10^{8}$ electrons	evaluate

Step 5 shows that a miniscule amount of charge − about a ten-thousandth of the charge you transfer to a balloon when you rub it on your shirt − is enough to balance the gravitational attraction between two one-kilogram masses separated by one meter. If you were concerned about whether adding the excess electrons would alter the mass of each sphere enough to require recalculating their gravitational attraction, you can compute that they add an insignificant mass, about 5×10^{-22} kg, to each sphere.

As an additional exercise, you can use Avogadro's number, and the atomic weight and atomic number of lead, to find the total number of electrons in an uncharged kilogram of lead. This calculation is not shown, but the total number is 2.39×10^{26} electrons. This means that the excess electrons constitute about 10^{-16} percent of the electrons in the sphere.

23.11 - Interactive checkpoint: electric vs. gravitational force



Compute the ratio of the electric to the gravitational force between the proton and electron in a hydrogen atom. Use the average distance between the two, which is called the Bohr radius and equals 5.29×10^{-11} m.

Answer:

$F_E / F_G =$

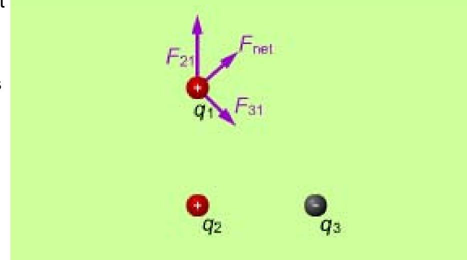
23.12 - Superposition of electrostatic forces

Electrostatic forces obey the *principle of superposition*. The forces caused by multiple charges can be added as vectors. For instance, consider the charges shown in Concept 1. To calculate the net force exerted by the other charges on the charge labeled q_1, the forces exerted by q_2 and q_3 on q_1 are individually calculated and then those two forces are added as vectors. In the next section, we solve a sample problem involving charges that requires the use of this principle.

You must be careful about the directions of electrostatic forces, especially when combining forces that may point in opposite directions. The location and signs of the charges determine the direction of the forces.

For example, consider another set of charges, q_4, q_5, and q_6, with all the charges on a line, and q_6 the rightmost charge. Let's say charges q_4 and q_5 have opposite signs. Since both are on the same side of q_6, the forces they exert upon it will act in opposite directions. There will be cancellation and the net force will be less than the sum of the magnitudes of the individual forces. There is no "rocket science" here. Just be careful to consider the direction of forces before combining them.

concept 1

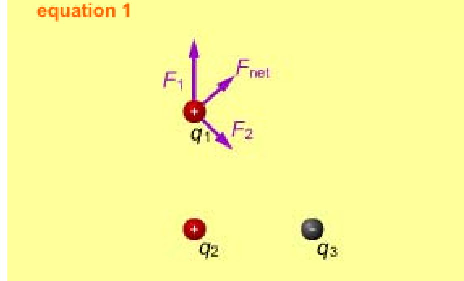


Electrostatic forces: vectors

Electrostatic forces are vector quantities

Net force = vector sum

equation 1

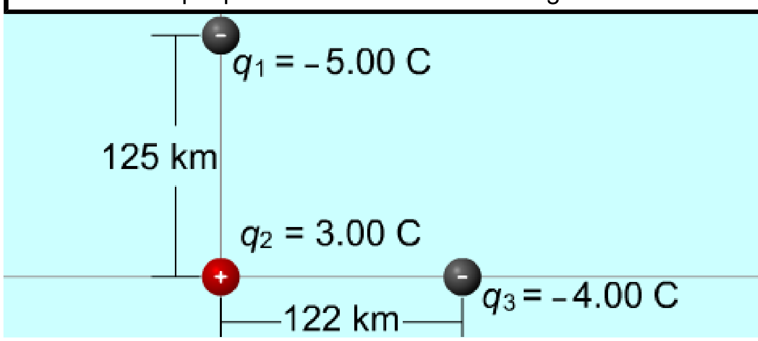


Electrostatic forces: vectors

Forces $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$

$$\mathbf{F}_{net} = \sum_{m=1}^{n} \mathbf{F}_m = \mathbf{F}_1 + \mathbf{F}_2 + \ldots + \mathbf{F}_n$$

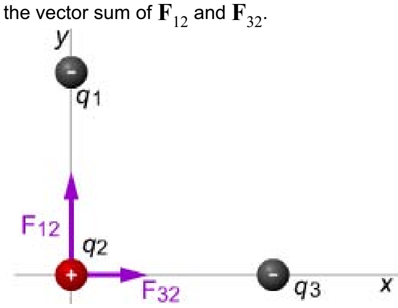
23.13 - Sample problem: net force on a charge



Find the magnitude and direction of the net force exerted by the charged spheres q_1 and q_3 on sphere q_2.

Diagram

The net force on q_2 is the vector sum of $\mathbf{F}_{12}$ and $\mathbf{F}_{32}$.



Variables

Coulomb's constant	$k = 8.99\times10^9\ \text{N·m}^2/\text{C}^2$	
magnitude of charge	q	
distance between charges	r	
	x component	y component
force of q_1 on q_2	$F_{12,x} = 0\ \text{N}$	$F_{12,y}$
force of q_3 on q_2	$F_{32,x}$	$F_{32,y} = 0\ \text{N}$
net force on q_2	F_{netx}	F_{nety}

In the table it is indicated that the x component of $\mathbf{F}_{12}$ is zero, since this vector points straight up. Similarly $\mathbf{F}_{32}$ points to the right, so its y component is zero. For the sake of brevity, we have not listed the individual values of all the charges and distances shown in the diagram above.

What is the strategy?

1. Find the x and y components of the force $\mathbf{F}_{12}$ exerted by sphere 1 on sphere 2.
2. Find the x and y components of the force $\mathbf{F}_{32}$ exerted by sphere 3 on sphere 2.
3. Add the vectors found above, component-by-component. Convert the resulting vector to polar notation to find its magnitude and direction.

Equations

We will use Coulomb's law

$$F = k\frac{|q_1||q_2|}{r^2}$$

In so doing, we may assume that all the charge of a charged sphere is concentrated in a single point at its center.

The magnitude and direction of the vector $\mathbf{A} = (a, b)$ are

$$\begin{cases} A = \sqrt{a^2+b^2} \\ \theta = \arctan\left(\dfrac{b}{a}\right) \end{cases}$$

Step-by-step solution

We use Coulomb's law to find the components of $\mathbf{F}_{12}$.

Step	Reason
1. $F = k\dfrac{\lvert q_1\rvert\lvert q_2\rvert}{r^2}$	Coulomb's law
2. $F_{12} = k\dfrac{\lvert -5.00\ \text{C}\rvert\cdot\lvert 3.00\ \text{C}\rvert}{(125{,}000\ \text{m})^2}$ $F_{12} = 8.63\ \text{N}$	evaluate
3. $\mathbf{F}_{12} = (F_{12,x}, F_{12,y}) = (0.00\ \text{N}, 8.63\ \text{N})$	components of $\mathbf{F}_{12}$

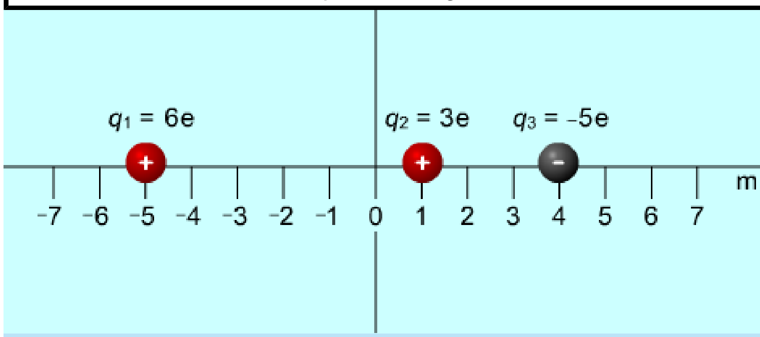
We use Coulomb's law to find the components of $\mathbf{F}_{32}$.

Step	Reason
4. $F = k\dfrac{\lvert q_3\rvert\lvert q_2\rvert}{r^2}$	Coulomb's law
5. $F_{32} = k\dfrac{\lvert -4.00\ \text{C}\rvert\cdot\lvert 3.00\ \text{C}\rvert}{(122{,}000\ \text{m})^2}$ $F_{32} = 7.25\ \text{N}$	evaluate
6. $\mathbf{F}_{32} = (F_{32,x}, F_{32,y}) = (7.25\ \text{N}, 0.00\ \text{N})$	components of $\mathbf{F}_{32}$

Finally, we add $\mathbf{F}_{12}$ and $\mathbf{F}_{32}$ as vectors, component by component, and find the magnitude and direction of $\mathbf{F}_{net}$.

Step	Reason
7. $\mathbf{F}_{net} = (F_{netx}, F_{nety}) = (7.25\ \text{N}, 8.63\ \text{N})$	vector addition
8. $F_{net} = \sqrt{F_{netx}^2 + F_{nety}^2}$ $F_{net} = \sqrt{(7.25)^2 + (8.63)^2}$ $F_{net} = 11.3\ \text{N}$	magnitude of vector
9. $\theta = \arctan\left(\frac{F_{nety}}{F_{netx}}\right)$ $\theta = \arctan\left(\frac{8.63}{7.25}\right)$ $\theta = 50.0°$	direction of vector

23.14 - Interactive checkpoint: charges on a line

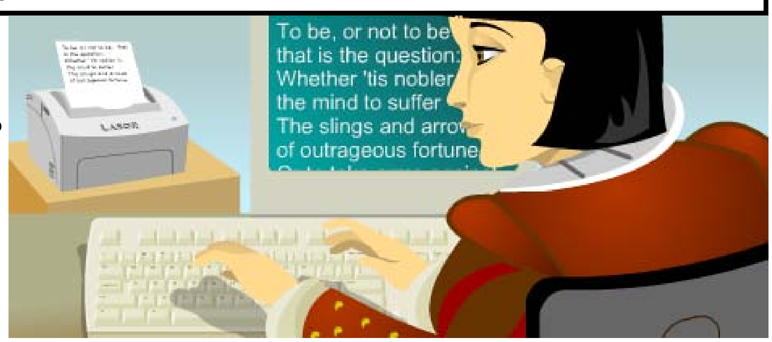


Three charges are arranged on a line as shown. Compute the net force exerted on the middle charge (charge 2) by the other two charges.

Answer:

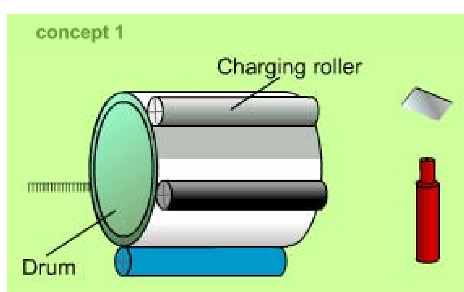
$\Sigma F =$ ______ N

23.15 - Physics at work: laser printers



When you print a document created on your computer, you may use a laser printer. Laser printers use the electrostatic force as a crucial part of the process of transferring an image from the computer to a sheet of paper. The process is known as *xerography* or *electrophotography*.

A key component of the laser printer is a rotating metal cylinder or drum. This drum is coated with a light-sensitive material called an organic photoconductor, a carbon-based compound whose electric properties change when it is exposed to light. In the dark it is an insulator, and electric charges cannot move through it. When it is exposed to light, it becomes a conductor and charges can flow freely. The organic photoconductor layer is on the outside surface of the drum. Inside it is a hollow metal cylinder connected to a ground, which allows any charge trapped in a portion of the photoconductor to drain away if that portion becomes conductive.



Charging
Photosensitive drum negatively charged

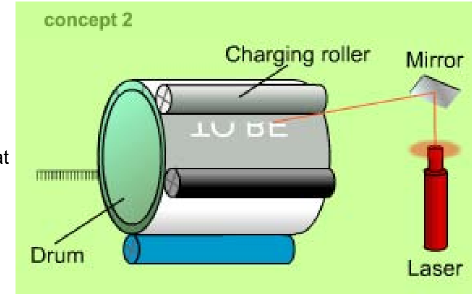
The printing process begins with a charging step, shown on the right, where the drum is given a uniform negative charge by bringing it into contact with a charged roller. Since the drum is shielded from light in the interior of the printer, the photoconductive layer acts as an insulator, trapping negative charges on its surface.

In the imaging step, shown next on the right, a laser controlled by signals from the computer directs light at the surface of the drum in a pattern corresponding to the image to be printed. Areas of the drum exposed to the laser light become conductive, allowing charge to escape from the drum surface to the ground. In this way the laser draws an electrostatic image of the document on the surface of the drum. Areas of the drum struck by the laser will be electrically neutral while the unexposed areas will retain their negative charge.

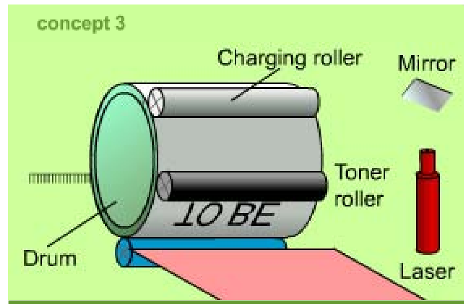
It is interesting to note that the laser itself is stationary. Its beam is projected onto the drum by a series of movable lenses and mirrors controlled by the printer's microprocessor, using data from your computer.

The next step is the development step. As the drum continues to rotate, it is brought close to a container or roller furnished with a toner, consisting of fine particles of ink that have a negative charge. The printer uses small pulses of electricity to eject the toner onto the drum surface. The negatively charged toner particles are repelled from the negatively charged unlit regions of the drum but cling to the neutral areas that were struck by the laser. The drum surface now holds toner particles in the pattern of the image to be printed.

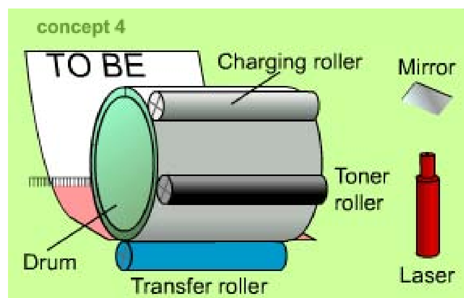
Then the image is transferred to paper. A sheet of paper is given a positive charge and pressed against the drum. The negatively charged toner on the drum is attracted to the positively charged paper. The toner is permanently fused to the paper by a heated roller that melts the ink particles into the paper fiber. A final step, not shown on the right, prepares the drum for the next image by flashing it with light, causing the complete discharge of all charged areas on the drum.



Imaging
Pattern of laser light strikes drum
Exposed areas lose charge



Development
Toner sticks to neutral exposed areas



Transfer
Positively charged paper captures toner

23.16 - Interactive summary problem: proton golf

Above and to the right, you see the 24th century version of golf. Protons in the putter cause the proton golf ball to move away. Coulomb's law is well known, well loved and well used.

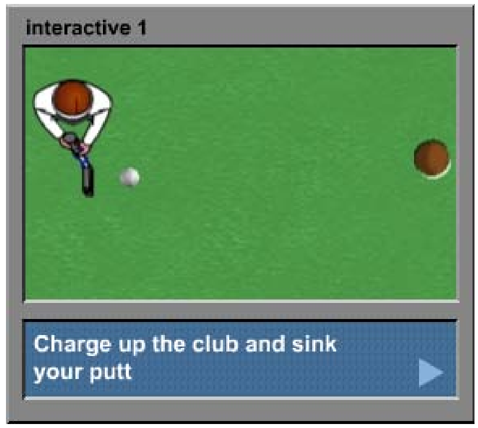
To sink the putt in the first game, your putter must supply an initial force of 1.96×10^{-25} newtons. This will cause the ball to be rolling slowly as it reaches the hole, overcoming the force of friction due to the grass. If you apply too much force, the ball will fly over the hole. The ball is free to move when you press PUTT. In this game, you cannot move the putter, and you get only one stroke, but you can play again by pressing RESET.

Proton Golf Association members at play in the year 2316.

The ball is initially 0.200 meters away from the putter and has a charge of $+e$. What should the charge be in the putter? You set the amount of charge by specifying the number of protons.

After you calculate your answer, click on Interactive 1 to launch the simulation. You use the up and down arrows to set the number of protons in the putter. Select your value, press PUTT and the golf ball will roll toward the hole. If you need to review how to calculate the repulsive force between two positive charges, see the section on Coulomb's law.

The second game is like the game of golf you played at the beginning of this chapter. You can change both the charge and the position of the putter. You will almost certainly need several strokes to sink the ball. The challenge of the game is to do so in as few strokes as you can. Again, the grass supplies a force of friction that must be overcome. You may feel that you are now more familiar with charges and the way they behave. Grab your prodigious proton putter and give it another try!



23.17 - Interactive group problem: minesweeper

On the right, you see an ocean overlaid with a grid. Hidden beneath the water are one or two mines loaded with electric charge. How many are hidden depends on which of the games you choose to play. Your mission is to deploy detectors to determine the location of the mines. The detectors will tell you the direction and amount of force being exerted on them, providing crucial information as to the location of the hidden mines.

To play the game, click on any of the graphics to the right. Each graphic includes text that describes the type of game (some are for one player, some are for two, and some have one hidden mine, while others have two).

This US Navy minesweeper has detectors for discovering hidden mines. Can you find the hidden mines in this section's interactive simulations?

Once you have launched a simulation, you can place the detector, which is a test charge, anywhere you like on the grid by dragging it from the control panel. Then click DEPLOY. If your detector is on exactly the same square as a mine charge, the charge will be revealed. The counter at the bottom of the game keeps track of the number of undetected mines.

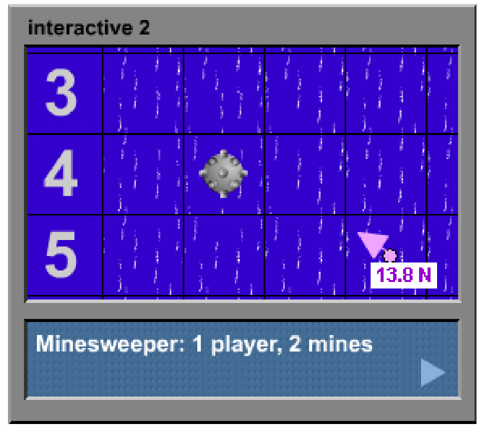
If you do not place the detector on the same square as a mine charge, you will see a vector. The vector tells you the direction and magnitude of the force that the mine charge exerts on your test charge. Most of the vectors you see will be drawn with lengths proportional to their actual magnitudes, but not the shortest ones. Those we draw a little longer so that you can see what direction they are pointing in.

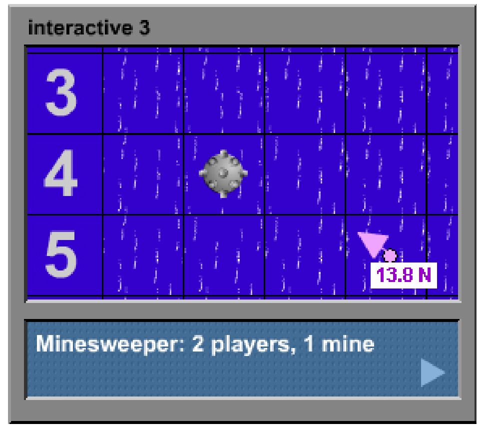
In all the games, the test charge in the detector has a magnitude of +0.005 00 coulombs. In the one-mine, one-player game, the mine charge is −0.010 0 C. Each square of the map grid is 50.0 meters on a side, and the mine charges and detectors are both placed at the centers of squares. When there is just one hidden mine of known charge, then with **one** test charge − and some calculation − you should be able to determine the location of the mine charge. Using a couple of detectors may make it easier!

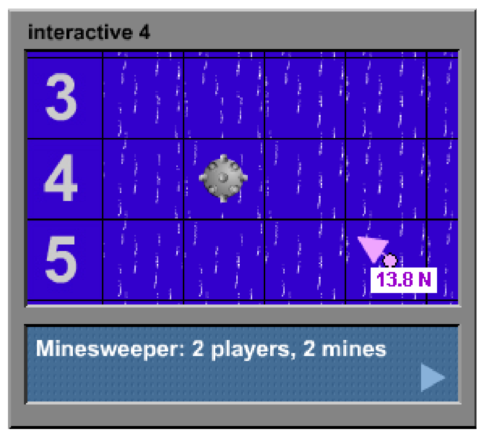
The two-mine game is harder. The force vector on any test charge is the vector sum of the forces from both hidden mines. There are two mines each with a charge of −0.010 0 C concealed, so it takes a little more experimenting and pondering to determine where they are.

There are also a couple two-player versions of the game. In both, you hide a charged mine or mines **on your opponent's board** (opposite from where you drag out your mines), and your opponent does the same on yours. (Look away when your opponent hides hers.)

You both get to choose from a set of varied charges. In one version of the game, only one charged mine is hidden. In the other, two charged mines are hidden. You drag detectors to locations on your own board, trying to find the mine(s) hidden there. Whoever discovers all the mines first, wins. All the games rely on the material in the sections on electrostatic force and Coulomb's law.







23.18 - Gotchas

A neutral object that gains electrons is negatively charged. Yes, a neutral object that **gains** electrons becomes **negative**; a neutral object that **loses** electrons becomes **positive**.

An object has a net charge of negative 10 coulombs. How many electrons does it contain? Do not bother trying to calculate a value. You could calculate how many **excess** electrons the object contains − how many more electrons than protons − with the given information. But unless someone tells you how many electrons the object contained when it was neutral, you cannot answer the question. The point here is: Charge refers to the number of excess electrons. A neutral object has electrons, too, but they are balanced by an equal number of protons.

You use the number of excess protons and electrons in Coulomb's law. No, the charges q_1 and q_2 in Coulomb's law are measured in (what else?) coulombs. If you are given the number of excess electrons or protons in a problem, you must determine the electric charge in coulombs.

I calculated a negative force from Coulomb's law. Then you erred. The **amount** (magnitude) of the force is calculated by multiplying the absolute values of the charges, so it will always be positive. The **direction** of the force will vary by sign: attraction when the signs are opposite, repulsion when they are the same.

23.19 - Summary

Electric charge is a property of matter. It occurs in positive and negative forms. One unit of charge is e. A proton has a charge of $+e$ and an electron has a charge of $-e$. Charge, represented by q, is measured in coulombs (C). The elementary charge e equals 1.602 18×10^{-19} C.

An ordinary object is charged when it has an imbalance of protons and electrons.

Charge is always conserved. Though charges may be transferred from object to object, charge cannot be created or destroyed, and the net charge of an isolated system will remain the same.

Equations

Coulomb's law

$$F = k\frac{|q_1||q_2|}{r^2}$$

Electrons flow more freely in some objects than in others. Conductors allow electrons to move relatively easily, while insulators do not. A ground can drain away any excess charge from a conducting object. The most common ground is literally the ground: Earth.

Charged particles exert an electrostatic force on each other. Unlike gravity, which is always attractive, the electrostatic force can be either attractive or repulsive. Opposite charges attract each other, while like charges repel.

Coulomb's Law describes the amount of the electrostatic force between two point charges. It is proportional to the product of the charges' magnitudes and inversely proportional to the square of the distance between them.

Conceptual Problems

C.1 Two scientists are locked in a bitter dispute about the charge on a particularly famous particle of dust. Maria claims it has a charge of 2.40×10^{-19} C. Richard disagrees; he thinks its charge is 3.20×10^{-19} C. Which scientist should you believe? Explain.

○ Maria ○ Richard

C.2 You are given an apple and an orange. The apple has a net charge of $+3\times10^{-17}$ C. The orange has a net charge of -3×10^{-17} C. With only this information, can you determine which one has more total electrons? Explain.

○ Yes ○ No

C.3 In a process known as beta decay, a neutron in an unstable atomic nucleus becomes a proton, in the process ejecting an electron and an antineutrino. (a) Use conservation of charge to determine the charge of an antineutrino. (b) Sixty billion neutrinos (mostly from the Sun) pass through every square centimeter on Earth every second. They are hardly noticeable due to their negligible mass and weak interaction with matter. When a neutrino and an antineutrino collide, however, they annihilate each other and produce two (electrically neutral) gamma rays traveling in opposite directions. What is the charge of a neutrino?

(a) __________ C
(b) __________ C

C.4 The Sun generates most of its light through a series of nuclear reactions called the proton-proton chain. The overall process of these nuclear reactions can be summarized as $4\ {}_1\mathrm{H} \rightarrow {}_2\mathrm{He} + 2\nu + 2\gamma + 2?$ where ν is a small neutral particle called a neutrino, and γ is a photon (a particle of light), which is also neutral. The pre-subscripts on the element symbols H (hydrogen) and He (helium) indicate the number of protons in the nucleus of each element. What is the charge of the unknown entity "?" ?

○ $+2e$ ○ $+e$ ○ 0 ○ $-e$ ○ $-2e$

C.5 Under typical conditions in a safe home, classify the following materials as conductors or insulators: (a) rubber, (b) iron, (c) copper, and (d) wood.

(a) Rubber is a(n) i. conductor ii. insulator .
(b) Iron is a(n) i. conductor ii. insulator .
(c) Copper is a(n) i. conductor ii. insulator .
(d) Wood is a(n) i. conductor ii. insulator .

C.6 Vanessa has been walking around all day, and electrons that rubbed off from the carpet have caused her to acquire a net negative charge. Explain why her hair might stand on end.

C.7 A positively charged eraser is placed near the "0 cm" end of a 10 cm metal ruler. As a result of the induced charge effect, which end of the ruler becomes positively charged: the "0 cm" end, or the "10 cm" end?

○ the "0 cm" end
○ the "10 cm" end

C.8 You are locked in a rubber room and given a pair of rubber gloves along with a positively charged bar of gold (marked with a "+") and two electrically neutral bars of gold. You will be released if you can produce a negatively charged bar of gold, and you get to keep the gold. Explain how you might accomplish this.

C.9 Two unequally charged pellets are held apart at a fixed distance. The charge on one of the pellets is halved, and the charge on the other is doubled. How this will affect the electric force one pellet exerts on the other? Explain.

i. The force will quadruple.
ii. The force will double.
iii. The force does not change.
iv. The force will be halved.
v. The force will be quartered.

C.10 A can of root beer and a can of cola are given different amounts of charge. Consider the cans to be point charges. By what factor must their separation be changed so that the electric force on the root beer can is the same (a) if the charge on each of the cans is doubled, and (b) if only the charge on the root beer is doubled?

(a) Multiply original separation by
i. .500
ii. 1.000
iii. 1.414
iv. 2.000
v. 4.000

(b) Multiply original radius by
i. .500
ii. 1.000
iii. 1.414
iv. 2.000
v. 4.000

C.11 Consider an isolated system of three stationary nonzero charges. No other forces are present. (a) Is it possible that any of the charges is experiencing zero net force? (b) If your answer to part "a" is yes, give an example. If no, explain why not.

(a) ○ Yes ○ No
(b)

C.12 A professor pets her cat, which is initially neutral, as is she (even in an election year). Immediately afterwards, the professor reaches for a doorknob and feels a shock as electrons travel from the knob into her hand. Is the cat positively charged, negatively charged or electrically neutral?

i. Positively charged
ii. Negatively charged
iii. Electrically neutral

C.13 Two objects are seen to electrostatically repel each other. From this information, can you tell (a) whether they are positively or negatively charged or (b) whether the product of their charges is negative or positive?

(a) ○ Yes ○ No
(b) ○ Yes ○ No

C.14 A llama has a charge of 1.0 μC, while a distant planet has a charge of 2.0 coulombs. How does the magnitude of the electric force that the llama exerts on the planet relate to the magnitude of the electric force that the planet exerts on the llama? Is it the same, greater or less?

i. Less
ii. The same
iii. Greater

Section Problems

Section 0 - Introduction

0.1 Using the simulation in the first interactive problem in this section, answer the following questions. What is the direction of the force between (a) two particles whose charges are the same sign and (b) two particles of opposite sign? (c) What happens to the magnitude of the force when the distance between two particles is increased?

(a) i. Towards each other
ii. Away from each other

(b) i. Towards each other
ii. Away from each other

(c) i. Increases
ii. Stays the same
iii. Decreases

Section 1 - Electric charge

1.1 The nucleus of a helium atom contains two protons, two neutrons, and no electrons. Neutrons have no net charge. What is the charge of the nucleus?

__________ C

1.2 An electron has a mass of 9.11×10^{-31} kg. What is the charge of 1.00 grams of pure electrons?

__________ C

1.3 An initially neutral ball bearing is hooked up to a machine that transfers electrons onto it. Afterwards, the ball bearing is analyzed and its charge is -2.81×10^{-9} C. Find the mass increase due to the electron transfer. An electron has a mass of 9.11×10^{-31} kg.

__________ kg

Section 2 - Creating charged objects

2.1 Chlorine is used in the making of computers and blood bags, and is found in household bleach as well as in the skin of the Ecuadorian tree frog. (a) A chlorine molecule (Cl_2) has 34 protons and 34 electrons. What is its charge? (b) A chlorine ion (Cl^-) has 17 protons and 18 electrons. What is its charge?

(a) __________ C
(b) __________ C

2.2 The game of checkers has two players, black and red. Each player has 12 pieces. Suppose each red piece has a charge of + 2.10e−15 coulomb and each black piece has a charge of -5.00×10^{-16} C. Find the net charge of the following systems: (a) a black piece and a red piece, (b) two black pieces and a red piece, (c) all the pieces.

(a) __________ C
(b) __________ C
(c) __________ C

2.3 An iron arrowhead has an initial charge of 3.35e−6 C. How many electrons are required to give it a charge of −2.82 μC?

__________ electrons

2.4 A capacitor is a device having two electrodes that can be used to store electric charge. When a battery is attached to the capacitor, a positive charge moves to one electrode and an equal negative charge collects on the other. The *capacitance* of this system is defined as the ratio of the magnitude of the charge on either electrode to the voltage of the attached battery. Capacitance is measured in farads (coulombs per volt). A 1.5 volt D-cell battery is attached to a 2.2 picofarad capacitor. (a) Find the charge that appears on the positive electrode. (b) Find the charge that appears on the negative electrode. (c) Find the net charge of the entire capacitor.

(a) __________ C
(b) __________ C
(c) __________ C

2.5 A cloud that is about to unleash a lightning bolt has a charge of +24 coulombs. In stormy weather, powerful updrafts of air carry electrons from one place in the storm system to another. If the cloud started out electrically neutral, how many electrons were removed from it to give it this positive charge?

__________ electrons

Section 3 - Conservation of charge

3.1 In an experiment, a particle called a pion (π) is observed to decay into two other particles, a muon and a neutrino. The muon then decays into an electron and two more neutrinos. Neutrinos are electrically neutral. (a) What is the charge of a muon? (b) Pions come in three types: π^+ has a charge of $+1.60\times10^{-19}$, π has a charge of -1.60×10^{-19}, and π^0 is electrically neutral. What kind of pion could decay as described in this experiment?

(a) ○ $-2e$ ○ $-e$ ○ 0 ○ e ○ $2e$
(b) ○ π^+ ○ π ○ π^0

3.2 In a complex circuit, three colored wires (red, green, and blue) are joined together inside a mysterious black box. An ammeter, a device that measures the rate of charge flow (current), shows you that charge enters the box at 3 coulombs per second through the red wire and leaves the box at 2 coulombs per second through the green wire. There is no buildup of charge inside the box. After 10 seconds has elapsed, how much charge has flowed out of the box through the blue wire?

__________ coulombs

Section 5 - Conductors, insulators, and grounds

5.1 A nine-volt battery has two terminals that are 1.0 cm apart. The space between them is filled by air. How much charge flows between the two terminals in 1.0 minutes?

__________ C

Section 6 - Interactive problem: charged rods

6.1 Using the simulation in the first interactive problem in this section, produce a rod with a charge of +1.000 μC in two turns. Explain the steps you took to achieve this.

6.2 Using the simulation in the second interactive problem in this section, produce a rod with a charge of +1.000 μC in five turns. Explain the steps you took to achieve this.

Section 7 - Electrostatic force

7.1 Three indistinguishable balloons are given charges of − 1.1 μC, −2.5 μC, and +2.0 μC respectively. You are given two of them at random, and you observe that they repel each other. Find the total charge of the balloons you have been handed.

__________ C

7.2 Danny has a pile of four metallic marbles. Each marble has a charge of either −0.2 μC or +0.4 μC. He observes that red attracts blue, blue attracts green, and green attracts black. Red and black are brought in contact with each other so that they have the same charge. Afterwards, Danny observes that red now repels blue. For each color of marble, use the radio buttons below to tell the **initial** charge of the marble.

Red ○ +0.4 μC ○ −0.2 μC
Blue ○ +0.4 μC ○ −0.2 μC
Green ○ +0.4 μC ○ −0.2 μC
Black ○ +0.4 μC ○ −0.2 μC

Section 8 - Inducing an electric charge

8.1 A pack of 100-dollar bills is charged so that each bill has a charge of −0.0100 μC. The pack is suspended by an insulating thread inside of a neutral safe made out of a conducting metal. No charge can flow between the bills and the safe, or the safe and its surroundings. (a) Explain how the induced charge effect will change the distribution of charges on the safe. (b) A burglar knows that there is no net electrical charge on the safe. If the charge on the outer surface has magnitude −4.00 μC, what is the charge on the inner surface of the safe? (c) How much money is in the safe?

(a)
(b) __________ μC
(c) __________ dollars

Section 9 - Coulomb's law: calculating electrostatic forces

9.1 Two grapes are given equal charges and held apart at a distance of 1.3 m. They experience a repulsive force of 2.2 N. Find the magnitude of the charge on each grape.

__________ C

9.2 Two 2.0 kg plastic garbage cans are sitting 2.6 meters apart on a sticky classroom floor (coefficient of static friction μ_s = 0.40). They are not moving. If the first one has a charge of 10 μC, find (a) the most negative possible charge and (b) the most positive possible charge for the other garbage can. Assume that the charges can be represented as point charges located at the cans' centers.

(a) __________ C
(b) __________ C

9.3 Two positively charged skaters, Stacy and Bob, are traveling straight towards each other on frictionless ice. Consider them to be point charges. When they are 5.0 m apart, Stacy feels an attractive electric force of 2.2 N towards Bob. (a) What is the magnitude of the attractive force that Bob feels from Stacy? (b) The magnitude of Stacy's charge is twice as much as Bob's. What is the magnitude of Bob's charge? (c) Bob has twice the mass that Stacy does. At the instant that they are 5.0 m apart, from Bob's reference frame, Stacy appears to be moving towards him with an acceleration of 7.3×10^{-2} m/s^2. Find Stacy's mass.

(a) __________ N
(b) __________ C
(c) __________ kg

Section 10 - Sample problem: electric vs. gravitational force

10.1 The single electron and the single proton in a hydrogen atom in its lowest energy state are separated by a tiny distance called the Bohr radius, 5.29×10^{-11} m. Both particles carry a charge of magnitude e = 1.60×10^{-19} C. (a) Is the force between them attractive or repulsive? (b) What is the magnitude of the force?

(a) i. Attractive
ii. Repulsive
(b) __________ N

10.2 Two steel juggling balls each carry a charge of 2.75 μC. There is a repulsive force between them of 1.45 N. What is the distance between the centers of the two balls?

__________ m

10.3 Two balls of radius 2.00 mm have a separation between their centers of 5.33 cm. The same electric charge is placed on both balls so that there is a repulsive force between them of 2.75 N. Assume that the charge is uniformly distributed over each ball. What is the magnitude of the charge?

__________ C

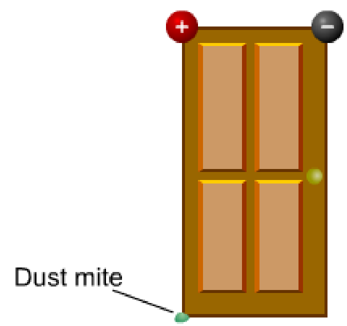
Section 12 - Superposition of electrostatic forces

12.1 Goldilocks is visiting the Three Bears' house and she notices three bowls of porridge arranged on the table. Papa Bear's porridge has a positive charge of +2.0 μC and is too hot. Mama Bear's porridge has a charge of −2.0 μC and is too cold. Baby Bear's porridge is just right and has a charge of +1.0 μC. If the bowls are in an equilateral triangle (with sides 0.50 m) what is the magnitude of the net force on Baby Bear's bowl?

__________ N

12.2 A doorframe is twice as tall as it is wide. There is a positive charge on the top left corner and an equal but negative charge in the top right corner. What is the direction of the electric force due to these charges on a negatively charged dust mite in the bottom left corner of the doorframe?

__________ °



12.3 Particles P_1 and P_2 are located in three dimensional space at the points (2.00, −3.50, 1.75) mm and (−3.50, 2.25, −2.00) mm. These particles carry charges of +3.00 μC and +4.50 μC respectively. What is the magnitude of the net force they exert on another particle P_3, with charge −3.50 μC, located at the origin?

__________ N

12.4 A positively charged ant is crawling through a plastic drinking straw with non-zero charges on both ends. One end is located at x = 0 and the other is located at x = 10 cm. The ant experiences zero net electric force at x = 3.0 cm. (a) Can you say whether the ends are positively or negatively charged? (b) Which end has a charge with greater magnitude? (c) Find the ratio of the charge at the 0 cm end to the charge at the 10 cm end. Express your answer as a decimal number.

(a) ○ Yes ○ No

(b) ○ The 0 cm end ○ The 10 cm end

(c) __________

12.5 Two positively charged objects and two negatively charged objects are kept in the corners (NE, NW, SW, and SE) of a square dorm room. The charge on each object has the same magnitude, *Q*, and no more than one charge can be kept in a corner. Describe the configuration of charges if the net force on a positive test charge in the center of the room (a) is zero and the charge in the NE corner is +*Q*, (b) points north, and (c) points west.

(a) NW corner i. +Q
ii. -Q

SW corner i. +Q
ii. -Q

SE corner i. +Q
ii. -Q

(b) NE corner i. +Q
ii. -Q

NW corner i. +Q
ii. -Q

SW corner i. +Q
ii. -Q

SE corner i. +Q
ii. -Q

(c) NE corner i. +Q
ii. -Q

NW corner i. +Q
ii. -Q

SW corner i. +Q
ii. -Q

SE corner i. +Q
ii. -Q

12.6 A bicycle wheel with radius 35 cm has 73 charges (1.0 μC each) spaced at regular intervals along its edge. A flea with charge −0.010 μC lands in the center of the wheel. (a) What is the net electric force on the flea? (b) One of the charges is removed. Find the magnitude of the new net electric force on the flea.

(a) __________ N
(b) __________ N

12.7 A mad scientist is designing a trap for intruders that will lift them up into the air and hold them helpless. The device consists of an equilateral triangle, 10.0 meters to a side, embedded in his floor. When he flips a switch, the corners of the triangle will be charged equally by a generator and any negatively charged object above the center of the triangle will be lifted upwards by the electric force. A computer-controlled system of giant fans keeps the intruder from straying horizontally from the center of the triangle. While he is testing the system, his cat walks into the trap. She has a net charge of −1.00 nanocoulombs due to electrons that rubbed off from the carpet. The cat, which has a mass of 5.00 kg, begins to hover 3.00 meters up in the air. Find the charge on each corner of the triangle.

__________ C

12.8 Four particles with charges of +1.00 μC, −2.00 μC, +3.00 μC, and −4.00 μC are located at the points (0, 0), (2.00, 0) cm, (2.00, 4.00) cm, and (0, 4.00) cm, respectively. What are the magnitude and direction of the total force they exert on a particle of charge +5.00 μC located at (1.00, 2.00) cm?

__________ N, directed in the
i. positive x
ii. negative x
iii. positive y
iv. negative y
direction

12.9 Three particles P_1, P_2, and P_3 are located at the points (−2.00, −1.00) m, (0, 2.00) m, and (3.00, −1.00) m, respectively. P_1 has a charge of 5.35 μC, P_2 has a charge of 6.03 μC, and P_3 has a charge of −2.75 μC. What are the magnitude and direction of the net force these three particles exert on a fourth particle of charge 2.50 μC, located at the origin?

__________ N, directed __________ °
i. above
ii. below
the x -axis.

12.10 Three particles P_1, P_2, and P_3 are located at the points (−2.00, −1.00), (0, 2.00), and (3.00, −1.00), respectively. P_1 has a charge of 5.00 μC, but the charges of P_2 and P_3 are unknown. However, the three particles exert no net force on a charged particle that is placed at the origin. You are asked to find the unknown charges. (a) Use the fact that the net horizontal force on the particle at the origin is zero to find the unknown charge on P_3. (b) Then use the fact that the net vertical force on the particle at the origin is zero to find the unknown charge on P_2.

(a) Charge on P_3 is __________ μ C
(b) Charge on P_2 is __________ μ C

12.11 A six-sided die has a positive charge in the center of each face, and a negative charge embedded in the center. The charge on each face is proportional to the value on the face. The top side of the die is 6; the bottom is 1; the east is 4; the west is 3; the north face is 5; the south is 2. Find the direction of the electric force on the charge in the center. Express your answer as a vector of length one by stating its components in the up, east, and north directions.

The unit vector has components _________ in the up, _________ in the east, and _________ in the north directions.

Section 16 - Interactive summary problem: proton golf

16.1 Using the information given in the first interactive problem in this section, how many protons should be loaded onto the putter to sink the ball in the hole (without banking it off the wall)? Test your answer using the simulation.

_________ protons

Additional Problems

A.1 Hydrogen is a flammable, nontoxic, colorless, odorless, and tasteless gas. The U.S. produces 100 billion cubic feet per year of hydrogen for industry and for the space program. A hydrogen atom consists of an electron $(9.11\times10^{-31}\ \mathrm{kg})$ and a proton $(1.67\times10^{-27}\ \mathrm{kg})$. (a) What is the net charge of a hydrogen atom? (b) Find the ratio of the electric to the gravitational force between the electron and the proton $(e\ 1.60\times10^{-19}\ \mathrm{C})$. (c) Suppose we adjusted the mass of the proton so that the ratio is 1. What would be the new mass of the proton? (d) Instead, suppose we adjusted the elementary charge, e, so that the ratio is 1. What would be the new elementary charge?

(a) _________ C
(b) _________
(c) _________ kg
(d) _________ C

A.2 A thin plastic pipe (0.60 kg/m) pivots at the origin, and extends along the *x* axis from *x* = 0 to *x* = 3.0 m. There is a very light particle with 2.6 μC of charge plugging the pipe at *x* = 3.0 m. Another charged particle (−3.0 μC) is at *x* = 3.0 m, *y* = 1.0 m. Find (a) the magnitude of the initial torque on the pipe and (b) its initial angular acceleration. Assume that the *xy* plane is horizontal.

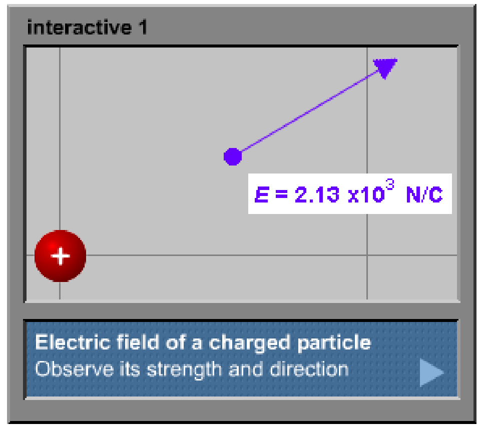
(a) _________ N · m
(b) _________ rad/s^2

A.3 Two 0.600 kg oppositely charged basketballs are following a clockwise circular path on a frictionless, freshly waxed basketball court. The balls are on opposite sides of the circle at all times, and are 10.0 m apart. Their charges cause the balls to continue on the circular path at a speed of 1.20 m/s. (a) Determine the product of the charges on the basketballs. (b) Now assume the charge on the positively charged ball is twice the magnitude of the negatively charged one. Determine the charge on the negative ball. (c) Determine the charge on the positive ball. (d) The same basketballs are now 5.00 m apart, but they are still moving in a circular path. Determine their speed. (e) One of the basketballs now has a mass of only 0.550 kg. Is it still possible for the two balls to hold each other so that they travel along identical circular paths? Explain your answer.

(a) _________ C^2
(b) _________ C
(c) _________ C
(d) _________ m/s
(e) ○ Yes ○ No

chapter 24 Electric Fields

24.0 - Introduction



An electric charge can exert force on another charge at a distance. No direct contact between the charges is required. The fact that the electrostatic, gravitational, and magnetic forces act at a distance puzzled early scientists who studied them. They speculated about the mechanism that would allow one body to push or pull on another without touching it. To explain how this could occur, the British scientist Michael Faraday (1791-1867) pioneered the concept of fields. As time has passed, this concept has assumed an increasingly important role in physics.

Today, physicists say that an electric field surrounds an electric charge, and that the electric field of one charge exerts a force on another charge. Investigating the fundamentals of fields is the central topic of this chapter.

To begin your study of fields, try the simulation to the right. Here, the charged particle you see is creating an electric field, represented by the symbol $\mathbf{E}$. The field is invisible but you can observe it using a field meter.

In essence, the field supplies a "road map" to calculate the force that will be exerted on any second charge that enters the region surrounding the charged particle you see. The stronger the field at a point in space, the greater the force that will be exerted on a given second charge when it is placed there. The field points in the same direction as the force that would be exerted on a positive test charge.

Locate your mouse pointer anywhere on the screen. When you click, you will see an electric field vector that points in the direction of the field, and a readout that tells you the strength of the electric field at that point.

The initial charge of the visible particle is 1.00 nC, but you can use the controller in the simulation to change this to other values, both positive and negative.

Initially, keep this particle's charge constant in the simulation, and try to answer the following questions: How does the field strength vary with the distance from the charge? Does it seem to increase linearly as you move closer, or do you see great increases in the field at points near the charge? You should see similarities between the electric field and the force that would be exerted on a second, positive charge.

Next, try changing the charge of the visible particle in the simulation. Does the amount of its charge affect its field strength at a given point? In what direction does the field of the charged particle point when it is positive? Does the field point in the same direction if you give the source particle a negative charge?

Using the simulation, you can experiment with some of the fundamentals of electric fields. In the rest of this chapter, you will continue your exploration of this topic.

24.1 - Electric fields

Electric field: An electric field describes the nature of the electric force that a charge will encounter at a given location. Fields provide the model for forces acting at a distance.

The muscles and heartbeats of fish generate telltale electric fields, perceptible to organs in the wide head of this hammerhead shark.

Electrically charged objects exert forces on other charged objects at a distance: The forces occur without direct contact. Charged objects exert forces on each other analogous to the gravitational forces exerted by the Earth and the Moon on one another. Electric, gravitational and magnetic forces all act at a distance. To describe their behavior, scientists have developed the concept of fields, which have become a fundamental tool for explaining the nature of the universe.

Charges establish electric fields around themselves. A small, positive *test charge* is often used to establish the nature of an electric field. A test charge is assumed to be weak enough that it does not alter the field being analyzed. In diagrams, we represent a test charge as a white sphere with a red plus sign.

You see in Concept 1 a test charge placed in a field, which in this case is generated by a negatively charged particle. The direction of the field is the direction of the force on the test charge. In the diagram, the field is represented as a vector and labeled with E, the symbol for an electric field. This field exists at every point whether or not a test charge is present.

As we explain the nature of electric fields, we first review the fundamentals of the electrostatic force. Coulomb's law states that the force between two charges is proportional to their magnitudes and inversely proportional to the square of the distance between them. Opposite charges attract one another, and like charges repel.

The strength of the field equals the amount of force divided by the magnitude of the test charge. The equation for calculating the field in this fashion is shown in Equation 1. Since an electric field is used to describe the nature of the force that a test charge experiences at a point in space, the field has a specific direction and magnitude at each point. It is a vector quantity.

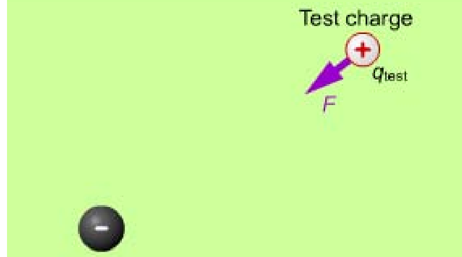
The strength of the field is independent of the test charge. Dividing out the charge cancels its effect in determining strength. For example, if the test charge were doubled in strength, then Coulomb's law states that the force would be twice as large, but after dividing by the doubled charge, the field strength is the same. Electric fields are measured in newtons per coulomb; there is no special name for this combination of units.

The stronger the field at a given point, the greater the force it will exert on any charge at that point. You can perhaps envision this by considering the nature of the gravitational field surrounding the Earth. The gravitational field is stronger near the surface of the Earth than it is at locations farther away. An object with a particular mass will experience more gravitational force closer to the Earth, where the field is stronger, than farther away, where the field is weaker. The gravitational field is stronger on the surface of Jupiter than it is on the Earth's surface. The stronger field means you weigh more on Jupiter because there is more gravitational force pulling on you.

In this section, we implicitly focused on fields produced by a single electric charge. This is an important case, and provides a concrete example of how fields arise. However, as you advance in your studies, you will often study fields and their effects by themselves, with less focus on their sources.

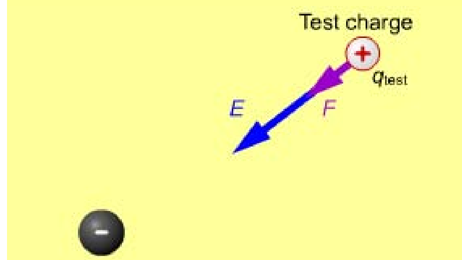
In the industrialized world, you are constantly surrounded by electric fields. You may experience electric fields ranging in strength from nearly zero up to 10 N/C due to electrical appliances and wiring as you walk around your house. Outside your house you may experience significantly greater electric fields. The field at ground level directly beneath a power transmission line is about 2000 N/C. This is enough to cause a fluorescent tube to light up if you hold it vertically in the field. Within atoms, electric fields are extremely large. The electric field due to a proton at the distance typical of an electron in a hydrogen atom is 5×10^{11} N/C.

concept 1



Electric fields
Charged object exerts force at a distance
Field surrounds charged object
Field equals force per unit charge

equation 1

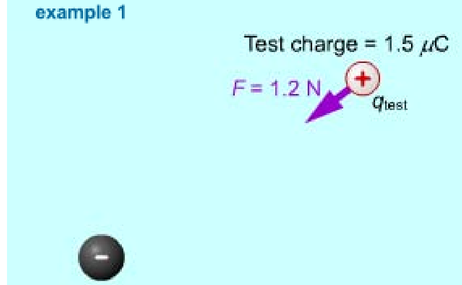


Electric field

$$E = F/q_{\text{test}}$$

E = strength of electric field
F = force on test charge
q_{test} = positive test charge
Direction of field: same as force
Units: newtons per coulomb (N/C)

example 1



What is the electric field at the location of the test charge?

$E = F/q_{\text{test}}$
$E = (1.2\text{ N}) / (1.5\times10^{-6}\text{ C})$
$E = 8.0\times10^{5}$ N/C
(same direction as force)

24.2 - Electric fields and Coulomb's law

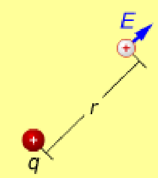
You can use Coulomb's law to calculate the strength of the field around a point charge. First, calculate the amount of force exerted by that charge on a (positive) test charge using Coulomb's law. Then divide by the amount of the test charge, because the strength of an electric field equals the amount of force divided by the charge used to measure the force.

We derive the equation for the field caused by a point charge as shown in Equation 1. We state the definition of the electric field, and then expand the expression for the force, using Coulomb's law. The quantity q_{test} appears in both the numerator and the denominator of the resulting fraction, so it cancels out. As the equation states, the strength of the field is proportional to the absolute value of q, the charge causing it, divided by the square of the distance from this charge.

An example problem is shown on the right, as well: You are asked to calculate the field strength at a point 0.110 m away from a charge of positive 10.5 C.

Remember: Even after you calculate the strength of the field, your task is not finished, because an electric field has both magnitude and direction. You must specify the field's direction, as we do in the example problem.

equation 1

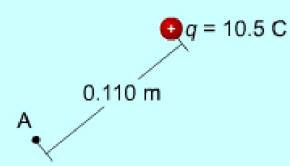


Electric field, point charge

$$E = \frac{F}{q_{test}} = \frac{k|q|q_{test}/r^2}{q_{test}} = \frac{k|q|}{r^2}$$

E = electric field
F = force
q = charge, q_{test} = test charge
k = Coulomb's constant
r = distance between charges

example 1



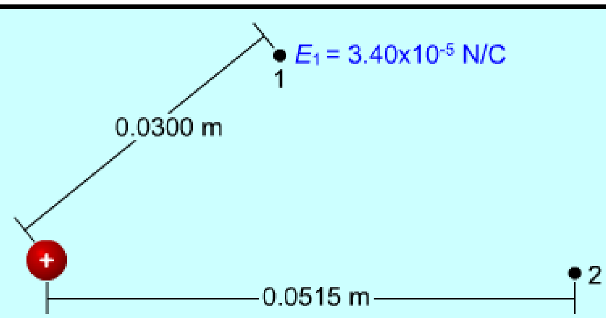
What are the magnitude and direction of the electric field at point A?

$$E = \frac{k|q|}{r^2}$$

$$E = \frac{(8.99 \times 10^9\ \mathrm{Nm^2/C^2})|10.5\ \mathrm{C}|}{(0.110\ \mathrm{m})^2}$$

$E = 7.80\times10^{12}$ N/C (away from charge)

24.3 - Interactive checkpoint: electric field



A point charge rests in an otherwise charge-free region. Locations 1 and 2 are 0.0300 m and 0.0515 m from the charge, respectively. If the electric field at location 1 is 3.40×10^{-5} N/C, what is the electric field at location 2?

Answer:

E_2 = ______ N/C

24.4 - Electric field diagrams

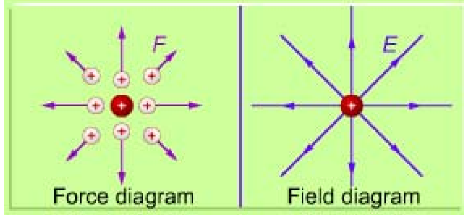
Electric field diagrams are used to illustrate the nature of an electric field. An electric field diagram consists of electric field lines, as you see illustrated in Concept 1.

The left-side diagram in Concept 1 shows the electric **force** acting on several test charges at various points around a central positive charge. The vectors representing the force on each test charge point away from the positive charge, indicating the direction of the force. The length of each vector is proportional to the amount of force on each charge.

The *field diagram,* shown to the right of the force diagram, is a common and efficient way to present similar information. The field lines have arrows pointing in the direction of the field vector at each point, which is the same as the direction of the force that the field would exert on a positive test charge at that point. However, the field lines do not use length to represent magnitude. Rather, field strength is indicated by how close the lines are to one another. Field strength is roughly proportional to the density of the lines.

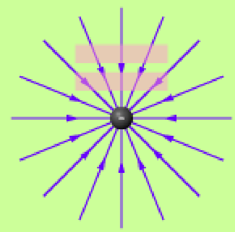
The illustration for Concept 2 shows how you can determine relative field strength from a field diagram. We have drawn two boxes at different distances from a negative charge. The field lines are closer together in the near box than they are in the far box, which means that the field is stronger in the region defined by the box closer to the negative charge.

concept 1



Electric field diagrams
Lines show force on positive test charge
Away from "+" charge, toward "–" charge

concept 2



Field line "density"
Proportional to field's magnitude

24.5 - Interactive problem: fields and forces

On the right are two simulations that allow you to explore electric field diagrams, and the relationship between electric fields and the electrostatic force.

In the first simulation, you explore the relationship between force and a field that is created by two unequal charges: a charge of +90.0 nC and a charge of −45.0 nC. In this simulation your task is to drag a positive test charge (+1.00 nC) into the field and observe the direction and magnitude of the force that the field exerts on it.

After experimenting with this simulation, consider the answers to three questions. First, is the force on the test charge greater where the field lines are closer together, or where they are farther apart? Second, what is the relationship between the direction of the force and the field lines? Hint: The word "tangent" must appear in your answer. Third, what…is the air-speed velocity of an unladen swallow? (Just kidding…But Monty Python fans know the consequences of a wrong answer.)

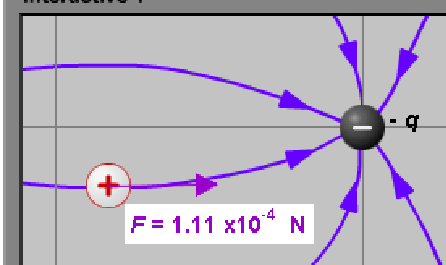
The *uniform field* in the second simulation is not caused by a single charge, or even by a small number of charges. The strength and direction of such a field are the same at all points. The field in this simulation points to the right, and its strength everywhere is 100 N/C.

In the simulation, you are given three small charges, having values $-q$, $+q$, and $+2q$. The negative charge is not, strictly speaking, a test charge, since test charges have to be positive. However, all three are so small that they do not interact with each other when they are placed in the uniform field.

Drag the three charges into the field, and observe the direction and magnitude of the force it causes on each one. How does the direction of the force exerted by the field differ for a positive and a negative charge? Does it exert a greater force on the $+2q$ charge or on the $+q$ charge? If so, by what factor is the force greater?

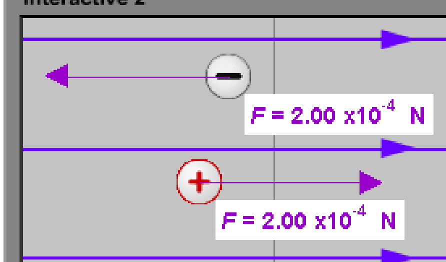
If you have trouble with the questions posed by either of these simulations, review the introduction to electric fields and electric field diagrams in previous sections.

interactive 1



Observe force on a test charge

interactive 2



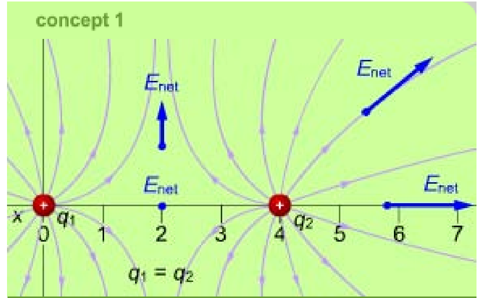
Observe forces on different charges

24.6 - Electric fields caused by multiple charges

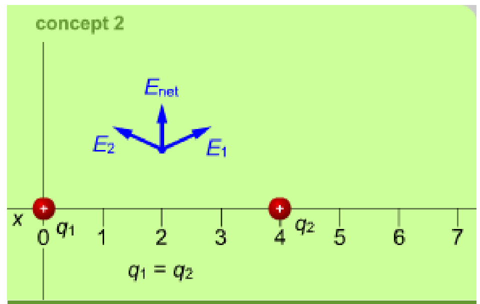
The individual fields from multiple charges can be combined to determine the overall field they create. The net electric field at any point is the vector sum of the individual fields due to each of the source charges. This additive property of electric fields is an example of the *principle of superposition*.

To calculate the overall field at a point, you start by calculating the field vector generated by each charge at that point. You then add the vectors together. The result is the net electric field at that point due to all the charges. In the illustration of Concept 1 you see a diagram of the combined field generated by two equal, positive charges. We also show the combined electric field vectors at several points in the field. In the upper right corner of the diagram, you can see that such a vector is tangent to a curving field line.

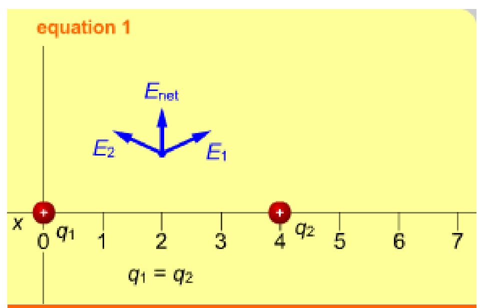
This method of combining fields is illustrated in detail in Concept 2 for one of the combined field vectors. The individual field vectors caused by the two identical positive charges at a location in space are depicted. We also draw the net field as the vector sum of the two individual fields. At the location shown, the x components of the contributing fields cancel out since they point in opposite directions and are of equal magnitudes. The y components of these two fields point in the same direction. The result is a combined field that points in the y direction, away from the two positive charges.



At each location in a field
E_{net} is sum of fields of each charge



Electric field of multiple charges
Calculate field vector due to each charge
Sum vectors to determine net field

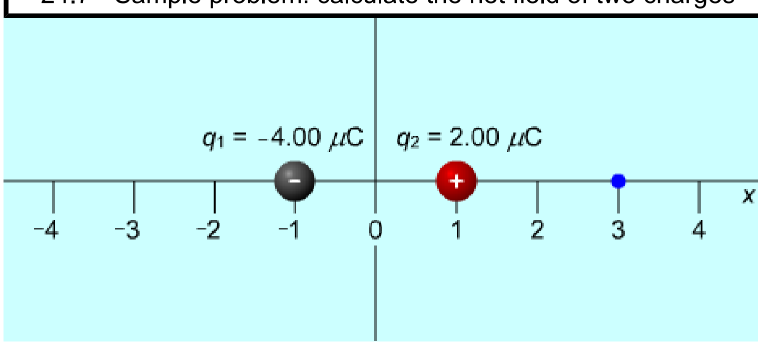


Electric field of multiple charges

$$\mathbf{E}_{net} = \sum_{i=1}^{n} \mathbf{E}_i$$

$\mathbf{E}_{net}$ = net electric field
$\mathbf{E}_i$ = electric field of a charge

24.7 - Sample problem: calculate the net field of two charges



Two charged particles are placed as shown at −1.00 m and +1.00 m on the x axis.

Calculate the combined field due to these charges at the indicated point, (3.00 m, 0.00 m).

Variables

Coulomb's constant	$k = 8.99\times10^9\ \mathrm{N{\cdot}m^2/C^2}$	
net electric field at point	$\mathbf{E}_{\text{net}}$	
	particle 1	particle 2
field due to charge	$\mathbf{E}_1$	$\mathbf{E}_2$
strength of field	E_1	E_2
direction of field	θ_1	θ_2
charge on particle	$q_1 = -4.00\ \mu\text{C}$	$q_2 = 2.00\ \mu\text{C}$
distance to particle	$r_1 = 4.00$ m	$r_2 = 2.00$ m

What is the strategy?

1. Find the x and y components of the fields $\mathbf{E}_1$ and $\mathbf{E}_2$ at the location stated in the problem.
2. Add the components to determine the net field $\mathbf{E}_{\text{net}}$.

Physics principles and equations

Opposite charges attract one another and like charges repel.

The equation for the strength of the electric field caused by a point charge is

$$E = \frac{k|q|}{r^2}$$

The x and y components of vector $\mathbf{A}$ with magnitude R and direction θ are

$$x = R\cos\theta$$

$$y = R\sin\theta$$

Step-by-step solution

We calculate the strength of the field caused by each charge using the formula above. We then determine its x component.

Step	Reason
1. $E_1 = \frac{k\lvert q_1\rvert}{r_1^2} = \frac{(8.99\times10^9\frac{\mathrm{N\cdot m^2}}{\mathrm{C^2}})(4.00\times10^{-6}\mathrm{C})}{(4.00\ \mathrm{m})^2}$ $E_1 = 2.25\times10^3$ N/C	equation for field created by point charge
2. $E_x = E_1\cos180° = -E_1$	direction of field due to sign of charge
3. $\mathbf{E}_1 = (-2250\ \text{N/C}, 0\ \text{N/C})$	express in rectangular notation

The formula for electric field strength and inspection of the diagram tell the magnitude and direction of $\mathbf{E}_2$. We write $\mathbf{E}_2$ in rectangular notation.

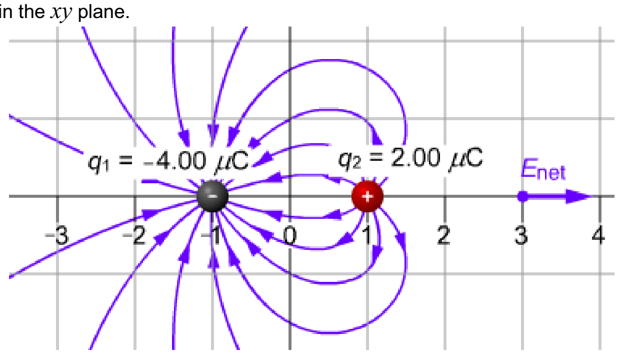
Step	Reason
4. $E_2 = \frac{k\lvert q_2\rvert}{r_2^2} = \frac{(8.99\times10^9\frac{\mathrm{N\cdot m^2}}{\mathrm{C^2}})(2.00\times10^{-6}\mathrm{C})}{(2.00\ \mathrm{m})^2}$ $E_2 = 4.50\times10^3$ N/C	equation for field created by point charge
5. $\theta_2 = 0°$	direction of field due to sign of charge
6. $\mathbf{E}_2 = (4500\ \text{N/C}, 0\ \text{N/C})$	express in rectangular notation

We have calculated the x and y components of both $\mathbf{E}_1$ and $\mathbf{E}_2$. We add the components of these two vectors to find $\mathbf{E}_{\text{net}}$ at the point (3.00 m, 0.00 m).

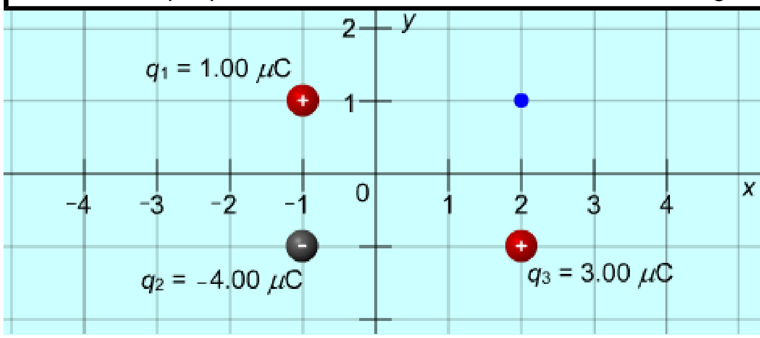
Step	Reason
7. $\mathbf{E}_{\text{net}} = \mathbf{E}_1 + \mathbf{E}_2$	principle of superposition
8. $\mathbf{E}_{\text{net}} = (2250\ \text{N/C}, 0\ \text{N/C})$	substitute equations 3 and 6 into equation 7

In the illustration below we display the vector $\mathbf{E}_{\text{net}}$ at the point (3.00 m, 0.00 m). For context, we also show an electric field diagram for $\mathbf{E}_{\text{net}}$.

The field diagram was drawn with the help of a computer program that automated the technique shown in this section to find $\mathbf{E}_{net}$ vectors for many points in the xy plane.



24.8 - Sample problem: calculate the net field of three charges



Three charged particles are placed as shown on the graph, exactly at the intersections of whole number grid lines.

Calculate the combined field due to these charges at the indicated point, (2.00 m, 1.00 m).

Variables

Coulomb's constant	$k = 8.99\times10^9$ N·m²/C²		
net electric field	$\mathbf{E}_{net}$		
	particle 1	particle 2	particle 3
field due to charge	$\mathbf{E}_1$	$\mathbf{E}_2$	$\mathbf{E}_3$
strength of field	E_1	E_2	E_3
direction of field	θ_1	θ_2	θ_3
charge on particle	$q_1 = 1.00\ \mu C$	$q_2 = -4.00\ \mu C$	$q_3 = 3.00\ \mu C$
distance to particle	r_1	r_2	r_3

What is the strategy?

1. Find the x and y components of the fields $\mathbf{E}_1$, $\mathbf{E}_2$ and $\mathbf{E}_3$ using the equation for the electric field caused by a point charge.
2. Using the principle of superposition, add the vectors found in the previous steps to find the net field $\mathbf{E}_{net}$.

Physics principles and equations

Opposite charges attract one another. Like charges repel.

The equation for the strength of the electric field caused by a point charge is

$$E = \frac{k|q|}{r^2}$$

The direction of the vector $\mathbf{A} = (a, b)$ is given by the angle

$$\theta = \arctan\left(\frac{b}{a}\right)$$

The x and y components of vector $\mathbf{A}$ with magnitude R and direction θ are

$x = R\cos\theta$

$y = R\sin\theta$

Mathematics principle

The Pythagorean theorem

$r^2 = a^2 + b^2$

Step-by-step solution

We use the formula for the electric field strength due to a point charge, and inspection of the diagram above, to determine the magnitude and direction of $\mathbf{E}_1$. We write $\mathbf{E}_1$ in rectangular notation.

Step	Reason
1. $E_1 = \frac{k\lvert q_1\rvert}{r_1^2} = \frac{(8.99\times 10^9 \frac{\mathrm{N\cdot m^2}}{\mathrm{C^2}})(1.00\times 10^{-6}\mathrm{C})}{(3.00\ \mathrm{m})^2}$ $E_1 = 1.00\times 10^3\ \mathrm{N/C}$	field strength due to point charge
2. $E_x = E_1 \cos 0° = E_1$	like charges repel, trigonometry
3. $\mathbf{E}_1 = (1000\ \mathrm{N/C},\ 0\ \mathrm{N/C})$	rectangular notation

The field caused by q_2 is more complex to calculate. We will use the Pythagorean theorem to calculate the distance. Then we use the same equation for field strength that we used above, and finally, some trigonometry to calculate the field components.

Step	Reason
4. $r_2 = \sqrt{(3.00\ \mathrm{m})^2 + (2.00\ \mathrm{m})^2}$	Pythagorean theorem
5. $E_2 = \frac{k\lvert q_2\rvert}{r_2^2} = \frac{(8.99\times 10^9 \frac{\mathrm{N\cdot m^2}}{\mathrm{C^2}})(4.00\times 10^{-6}\mathrm{C})}{(3.00\ \mathrm{m})^2 + (2.00\ \mathrm{m})^2}$ $E_2 = 2.77\times 10^3\ \mathrm{N/C}$	equation for field strength created by point charge
6. $\theta_2 = 180° + \arctan\left(\frac{2.00\ \mathrm{m}}{3.00\ \mathrm{m}}\right)$ $\theta_2 = 214°$	direction formula with q_2 negative
7. $\mathbf{E}_2 = (-2300\ \mathrm{N/C},\ -1550\ \mathrm{N/C})$	use trigonometry to calculate x, y components

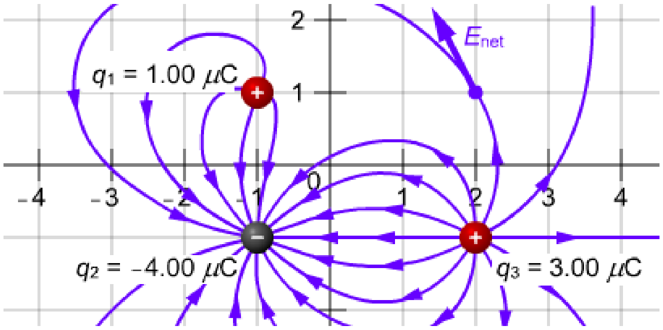
We follow a process identical to our first set of steps with charge q_1 here.

Step	Reason
8. $E_3 = \frac{k\lvert q_3\rvert}{r_3^2} = \frac{(8.99\times 10^9 \frac{\mathrm{N\cdot m^2}}{\mathrm{C^2}})(3.00\times 10^{-6}\mathrm{C})}{(2.00\ \mathrm{m})^2}$ $E_3 = 6.74\times 10^3\ \mathrm{N/C}$	field strength due to point charge
9. $E_y = E_3 \sin 90° = E_3$	like charges repel, trigonometry
10. $\mathbf{E}_3 = (0\ \mathrm{N/C},\ 6740\ \mathrm{N/C})$	rectangular notation

We have calculated the x and y components of $\mathbf{E}_1$, $\mathbf{E}_2$ and $\mathbf{E}_3$. We add the components of these three vectors to find $\mathbf{E}_{net}$ at the point (2.00 m, 1.00 m).

Step	Reason
11. $\mathbf{E}_{net} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3$	principle of superposition
12. $\mathbf{E}_{net} = (-1300\ \mathrm{N/C},\ 5190\ \mathrm{N/C})$	substitute equations 3, 7, and 10 into equation 11

In the illustration below we display the vector $\mathbf{E}_{net}$ at the point (2.00 m, 1.00 m). For context, we also show an electric field diagram for $\mathbf{E}_{net}$. The field diagram was drawn with the help of a computer program that automated the technique shown in this section to find $\mathbf{E}_{net}$ vectors for many points in the xy plane.



24.9 - Drawing field diagrams for multiple charges

Drawing the field diagram of a field generated by multiple charges requires the application of three rules. To explain them, we use the configuration of two charges shown in Concept 1. The negative charge is twice as strong as the positive charge.

Rule 1: The number of field lines in a diagram that emanate from a positive charge, or terminate at a negative charge, is proportional to the magnitude of the charge. In the diagram to the right, 8 lines emanate from q, and 16 terminate at $-2q$.

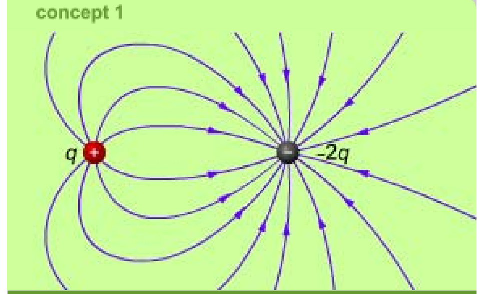
Rule 2: Other field lines may start from or extend to infinity. The "other" eight field lines that terminate at the negative charge, the ones that do not emanate from the positive charge, start at infinity. In other words, they enter from "outside" the diagram. If the positive charge were $+4q$, there would be more field lines emanating from it than terminating at the negative charge. The "extra" field lines would extend to infinity.

Rule 3: Field lines never cross. A positive test charge located at such an intersection would be "confused" as to what force it experienced, since the field would point in two directions. This rule reflects an important point: At any location, a field has a single direction and magnitude. This point is emphasized in the Concept 2 illustration.

These rules enable you to draw a correct field diagram. They provide a starting point for drawing a field diagram like the one shown in Concept 1. Considering the nature of the force experienced by a positive charge at various locations is required to flesh out the diagram.

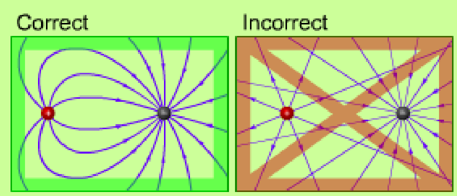
A last point to keep in mind: Electric field diagrams are commonly drawn in two dimensions, but electric fields exist in three dimensions. You usually see field diagrams on flat surfaces, like a computer monitor, a piece of paper or a blackboard. However, the field exists above and below the surface of the diagram as well.

concept 1



Field lines
Incident lines proportional to charge
Start at "+" charges, end at "- " charges
· Excess lines start or end at infinity

concept 2



Field lines
Do not cross

24.10 - Describing the force exerted by an external electric field

In this section, we shift our attention from the field caused **by** a charge to the force exerted **on** a charge by an external electric field. Such a field could be created by just one other charged particle, or it could be created by a much larger number of charged particles.

Fields are used in many everyday devices. For example, a field is used in some high-end inkjet printers to control the direction of charged ink droplets as they fly toward a sheet of paper. A charged camera flash can store energy in a field created by more than a trillion excess electrons. Given this huge number, determining the overall field by summing the individual fields created by each electron would be a bit tedious. Instead, placing a test charge in the field and observing the force that the field exerts on it can determine the nature of the field.

Polluting smokestacks. *Electrostatic precipitators* could electrically charge this particulate matter and use electric fields to remove it from the effluent.

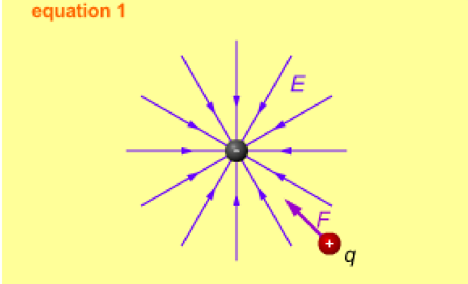
When a charged particle is in an electric field, the equation in Equation 1 can be used to describe the amount of force that the field exerts on the particle. This equation comes from the defining equation for the electric field, solved for the force.

A point charge creates a field that diminishes with the distance from the charge. This means the force it exerts on a test charge will vary by location. However, engineers are clever enough to create nearly *uniform fields*, fields whose strength and direction is the same at all locations within the field. For instance, there is a nearly uniform field in the center of the region between two large, oppositely charged flat plates separated by a small distance. In Example 1, you are asked to determine the force exerted on an electron by such a field.

Newton's second law enables you to determine the acceleration of the electron in this electric field and we do so in Example 2. The force exerted by the field is divided by the mass of an electron to determine the acceleration. The result is a large acceleration, having a magnitude of approximately 350,000 m/s^2.

This acceleration is caused by a relatively weak field (it is well below the maximum allowed for human exposure by government safety standards). At the acceleration stated, the speed of the electron would quickly approach that of light, and the acceleration would have to diminish due to the relativistic effects predicted by Albert Einstein.

equation 1

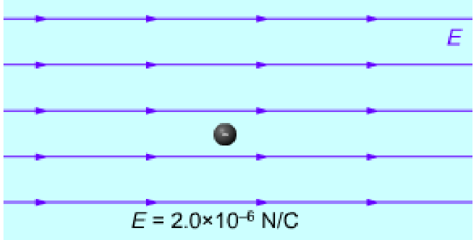


The force due to an electric field

$$\mathbf{F} = q\mathbf{E}$$

$\mathbf{F}$ = force
q = charge
$\mathbf{E}$ = electric field

example 1



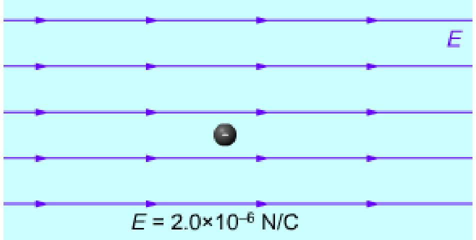
What is the force on the electron?

$\mathbf{F} = q\mathbf{E}$

$F = (-1.6\times10^{-19}\ \text{C})(2.0\times10^{-6}\ \text{N/C})$

$F = -3.2\times10^{-25}$ N (opposite to field)

example 2



What is the acceleration of the electron? Its mass is 9.1×10^{-31} kg.

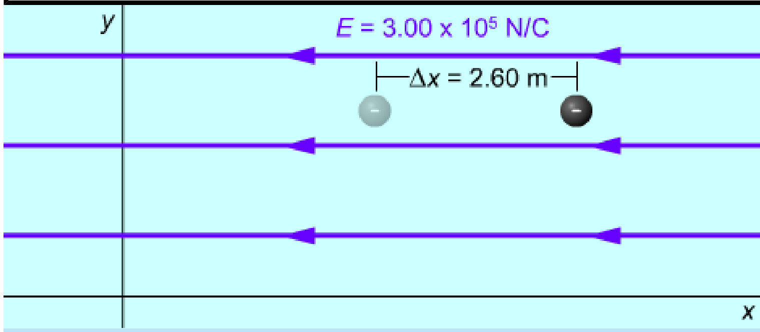
$a = F/m$

$m = m_e = 9.1\times10^{-31}$ kg

$a = (-3.2\times10^{-25}\ \text{N})/(9.1\times10^{-31}\ \text{kg})$

$a = -3.5\times10^{5}\ \text{m/s}^2$ (opposite to field)

24.11 - Interactive checkpoint: work on a charge

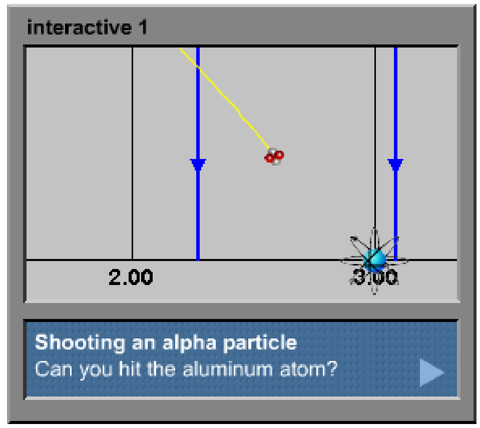


A uniform electric field of strength 3.00×10^5 N/C is oriented in the $-x$ direction. How much work is done by the field to move an electron 2.60 m in the $+x$ direction?

Answer:

$W =$ ______ J

24.12 - Interactive problem: the alpha cannon



This simulation asks you to recreate part of a famous experiment conducted by the 1935 Nobel Prize winners Frédéric Joliot and Irène Joliot-Curie (the daughter of Pierre and Marie Curie). You can learn more about their work and their prize at the Nobel Museum website.

Certain radioactive substances, such as radium, spontaneously emit various forms of radiation, including alpha particles. An alpha particle is a small charged particle consisting of two protons and two neutrons (neutrons have the same mass as protons but no electric charge). The Joliot-Curies used electric fields to trap and then accelerate these particles toward a target that consisted of aluminum atoms.

When an alpha particle strikes an aluminum atom, it fuses with the nucleus of the atom, a neutron is emitted, and an atom of radioactive phosphorus is the result. Research concerning the causes and nature of radioactivity won Nobel Prizes in both chemistry and physics for members of the Curie family.

In the simulation to the right, you can recreate a portion of one of their experiments. In this simulation you set the initial horizontal velocity of the alpha particle. There is a uniform electric field of 3.60×10^{-4} N/C directed downward. The electric charge of a proton is $+1.60\times10^{-19}$ C and its mass is 1.67×10^{-27} kg. This means the charge of the alpha particle is 3.20×10^{-19} C and its mass is 6.68×10^{-27} kg.

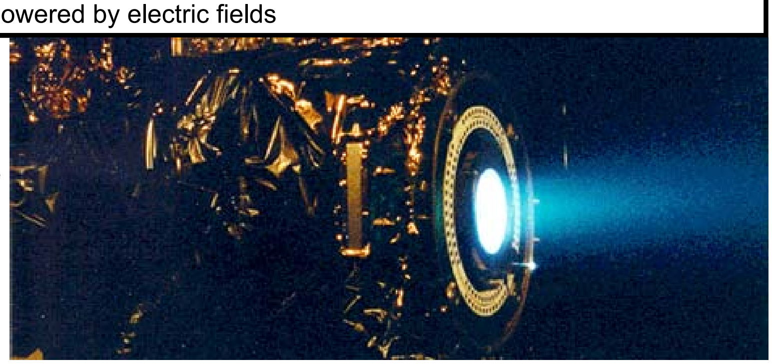
The simulation asks you to combine your knowledge of fields with concepts from the study of motion and vectors, specifically projectile motion.

Click on the graphic to the right to launch the simulation. Observe the vertical and horizontal distances of the cannon from the aluminum atom target, calculate an initial horizontal velocity to the nearest 0.1 m/s, and enter it in the space provided. The acceleration due to the electric field is great enough that you can ignore the acceleration of the alpha particle due to gravity.

When you press GO, the cannon will fire the particle with the initial velocity you specify. With the right calculations, you can make the alpha particle hit the atom, transmute it to phosphorus, and win your very own Nobel Prize (currently worth about a million dollars)!

The simulation does not run in real time, but is slowed by a factor of 1000. What takes place in one second by your watch is taking one millisecond in the simulation. If you have trouble hitting the atom, review the material in the previous section on the force exerted by an electric field, and the discussion of projectile motion in the chapter on motion in two and three dimensions.

24.13 - Physics at work: spacecraft powered by electric fields



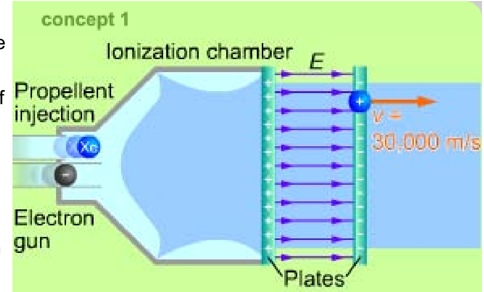
NASA's *Deep Space 1* ion drive undergoing testing at the Jet Propulsion Laboratory.

In 1998, NASA began testing in space a new type of rocket drive: the ion propulsion system. A conceptual diagram (not drawn to scale) of this drive is shown in Concept 1. Traditional rocket engines use chemical reactions to expel rocket fuel exhaust from the engine at high speed. If you have seen video footage of the rockets that carried astronauts to the Moon, you have seen a traditional rocket engine.

In contrast, the ion drive uses an electric field to accelerate charged gas particles, or ions. It accelerates these particles to extremely high velocities, velocities much greater than those achieved by chemical rocket exhausts. Why are higher velocities of interest? Recall the law of conservation of momentum. In this case, the “forward” momentum of the rest of the rocket must match the “backward” momentum of the rocket exhaust. Momentum is the product of mass and velocity. Since the ion drive accelerates particles to greater velocities than those of chemical drives, less fuel mass is necessary than with chemical systems to produce the same amount of momentum.

Less massive space probes are important because they require less energy to launch into space. In practice, a space probe is still launched by a powerful chemical rocket engine, but once it is relatively free of the Earth’s gravity, the probe relies on a less massive ion propulsion system to maneuver. (Ion systems supply less than one newton of force, a small fraction of the force required to launch a rocket from the Earth’s surface.) The system’s lesser mass also means it accelerates more for a given force.

The first step in the operation of this engine system is to bombard xenon gas with electrons. These fast electrons “knock off” an electron from each xenon atom, turning it into a positive ion. The ions are then inserted into an electric field created by two charged plates. The plates are less than a meter in diameter and are placed close together. The electric field between the plates accelerates the ions to a velocity of about 30,000 m/s (70,000 mi/h). The ions escape through small holes in the rear plate. (To prevent the ions from being attracted back into the engine, a beam of electrons is fired into the ion exhaust to neutralize it.)

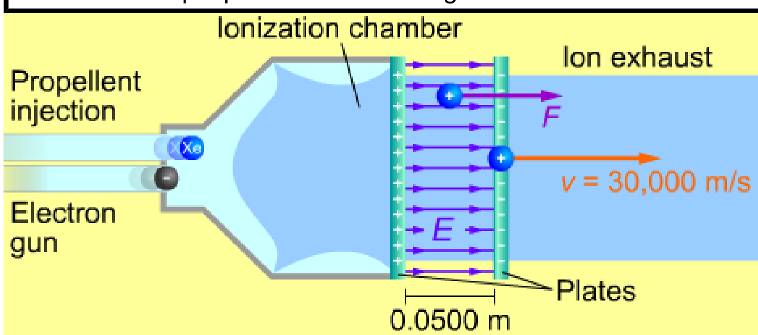


An electric-field rocket engine
Electron bombardment ionizes xenon gas
Field accelerates ions to 30,000 m/s
Low mass, high velocity system

NASA tested this engine system on the Deep Space 1 mission. Deep Space 1 was shut down on December 18, 2001, after successfully validating the performance of the ion propulsion system. The drive had provided more than 16,000 hours of thrust, using only 72 kg of xenon gas. This was less than 10% of the overall mass of the spacecraft. An equivalent chemical rocket drive would have needed 720 kg of fuel to generate the same amount of propulsion. This is equivalent to the difference between launching a person and launching a small car.

The ion propulsion system is now in use in satellites that orbit the Earth, where it is used to “tune” their orbits. Photovoltaic panels supply the energy required to ionize the gas and charge the plates.

24.14 - Sample problem: field strength of an ion drive



How strong is the electric field?

This problem asks you to be a rocket scientist. Xenon ions are expelled from the drive at a speed of 30,000 m/s. The ions have a charge of $+e$ and start at the left plate with an initial velocity that is effectively zero. They accelerate in a uniform field between the plates, which are separated by a distance of 0.0500 meters.

Variables

initial speed of ion	$v_i = 0$
final speed of ion	$v_f = 30{,}000$ m/s
distance between plates	$\Delta x = 0.0500$ m
acceleration of ion	a
mass of ion (= atomic mass)	$m = 2.20\times10^{-25}$ kg
electrostatic force on ion	F
charge of ion (= $+e$)	$q = +1.60\times10^{-19}$ C
strength of electric field	E

What is the strategy?

1. Use an equation from the study of linear motion to calculate the required acceleration of the ion. The acceleration is constant because the field is uniform.
2. Use Newton’s second law to determine the force.
3. Finally, use the equation for electric field that relates field, force and charge.

Physics principles and equations

We will use the kinematics equation

$$v_f^2 = v_i^2 + 2a\Delta x$$

Newton's second law

$$F = ma$$

The equation defining field strength is

$$E = F/q$$

Step-by-step solution

In the first steps, we find the acceleration of a xenon ion between the charged plates.

Step	Reason
1. $v_f^2 = v_i^2 + 2a\Delta x$	standard motion equation
2. $a = \dfrac{v_f^2 - v_i^2}{2\Delta x}$	solve for a
3. $a = \dfrac{(3.00\times 10^4 \text{ m/s})^2 - 0^2}{2(0.0500 \text{ m})}$ $a = 9.00\times 10^9 \text{ m/s}^2$	evaluate

Now we find the force necessary to give each ion the calculated acceleration.

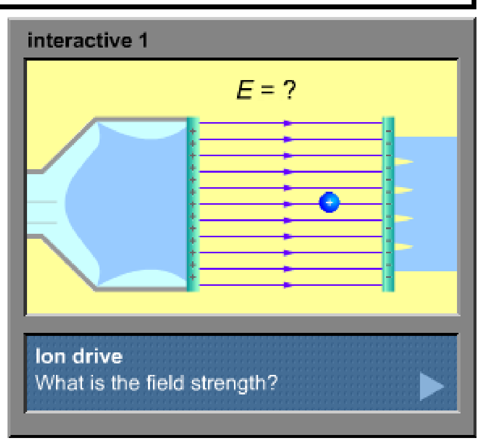
Step	Reason
4. $F = ma$	Newton's second law
5. $F = (2.20\times 10^{-25} \text{ kg})(9.00\times 10^9 \text{ m/s}^2)$ $F = 1.98\times 10^{-15}$ N	evaluate

Finally, we use the force to calculate the strength of the electric field.

Step	Reason
6. $E = F / q$	equation for field
7. $E = \dfrac{1.98\times 10^{-15} \text{ N}}{1.60\times 10^{-19} \text{ C}}$ $E = 12{,}400$ N/C	evaluate

The field strength required is 12,400 N/C. For points of comparison, the field around a charged balloon would be in the range of 1000 N/C, and the field around a laser printer drum is about 100,000 N/C.

24.15 - Interactive problem: tune the rocket's drive field



Xenon is a convenient substance to use in an ion rocket drive because it is fairly heavy, reacts with few other elements and is non-radioactive. In fact, it is the heaviest non-radioactive noble gas.

In the simulation to the right, we have changed the gas used in the engine from xenon to radioactive radon. (We did so to make some of the parameters in the problem different from those for xenon, we must admit, and not because some NASA engineer has changed her mind.)

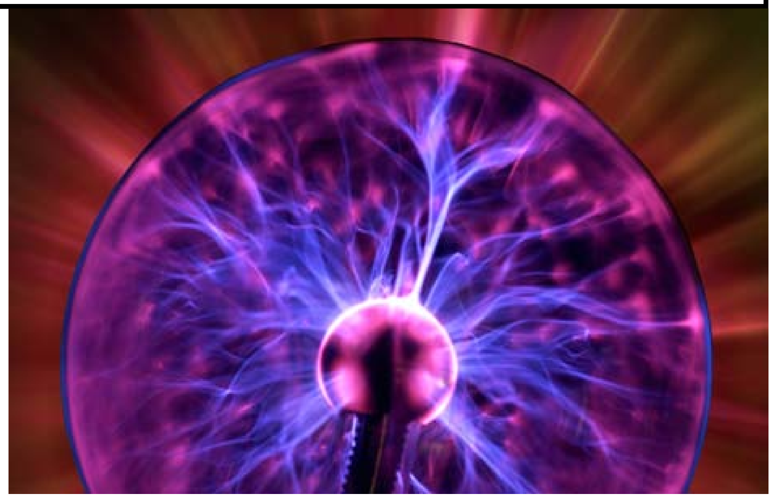
In the engine, an electron is stripped from the radon molecule, leaving it with a charge of $+e$. The positively charged radon ion has a charge of 1.60×10^{-19} C and a mass of 3.71×10^{-25} kg. A distance of 0.155 m separates the oppositely charged electric plates of the new drive. Since radon ions are more massive than xenon ions, an exhaust velocity of 17,700 m/s is adequate for the needs of the drive.

Your job is to calculate the uniform electric field strength necessary to accelerate the ions, from rest, to the desired exhaust velocity. You can test your calculation in the simulation to the right. Once you have computed the field strength (a positive value), enter it in the space provided, and then click on GO. You will be shown the exhaust velocity that your field gives to the radon ions.

Once you get the correct field strength for an exhaust velocity of 17,700 m/s, you will be a real rocket scientist. For more information about the calculations required to find the field strength, review the definition of the electric field, and the equations from the chapter on motion in one dimension.

24.16 - Conductors in electrostatic equilibrium

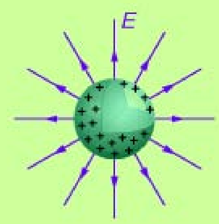
Electrostatic equilibrium: In an isolated conductor, excess charges quickly achieve a state where there is no net motion of charge.



Charge resides on the surface of the inner metallic sphere in this toy. An electric field extends radially outward, made visible by ionized gas.

An isolated, charged conducting object has several interesting and perhaps unexpected properties. (By isolated, we mean the object is attached neither to a source of what is called a potential difference, such as a battery, nor to a ground.) The excess charge on such an object will rapidly reach a state called electrostatic equilibrium, which means there is no further net motion of the charge. In other words, any excess charges in the object can be treated as stationary. Many experiments have confirmed the existence of electrostatic equilibrium and the properties of conductors described below.

concept 1



Electrostatic equilibrium
Excess charge moves to surface
Electric field perpendicular to surface
No field inside material of conductor

What are some properties of a charged conductor in electrostatic equilibrium? First, it creates an electric field that, just outside the conductor's surface, is perpendicular to the surface. You see this in the diagram for Concept 1, which shows a solid conducting sphere.

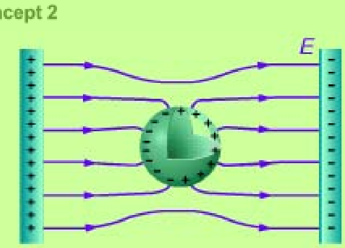
Why is the field perpendicular? This question can be answered by asking another question: What if the field were **not** perpendicular? If there were a component of the field parallel to the surface, this would mean a charged particle there would have a force exerted on it by the field, parallel to the surface. This force would cause the particle to move along the surface, and this motion would contradict the assumption of electrostatic equilibrium.

Second, all excess charge resides on the surface of a conducting object. One way to explain this property is to note that such a configuration allows like charges to maximize their distance from one another. If a charge remained in the middle of the conductor, it would not be maximizing its distance from its fellows.

Third, there is no electric field within the bulk of a conductor. Again, consider the conducting sphere in Concept 1. If there were a field within the sphere, it would cause movement of excess charges there, meaning the sphere would not be in electrostatic equilibrium.

(Gauss' law, another principle of electrostatics that you will learn about, can more economically explain the fact that there is no field inside a conductor.)

concept 2



Hollow conductor
Shields contents from external fields

In Concept 2, we show what occurs when a hollow conductor is placed in an external field. In this case, two charged plates cause the field. If the sphere were not present, the electric field would be uniform, and a field diagram would represent the field with equally spaced, horizontal lines.

When the sphere is placed in the field caused by the plates, an asymmetrical distribution of charge is induced on the outside of the sphere. You can see the charge distribution and how the sphere alters the external field in the diagram.

There is another important point illustrated in the diagram: There is no field within the hollow conductor. This effect is quite useful because it means a conducting shell will insulate its contents from electric fields. When you see electric or electronic circuits placed within a protective metal box, one reason is to shield them from nearby electric fields that might distort the circuitry's operation. The shielding effect occurs only with conductors. It cannot occur with insulators because charge is not free to move within such substances.

24.17 - Electric dipoles

Electric dipole: Two point charges with equal strength but opposite signs separated by a fixed distance.

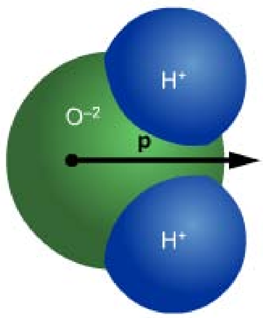
A configuration of two charges that are equal in strength but opposite in sign is important enough to merit its own name: the electric dipole. The two charges are at a fixed distance from each other.

The diagram in Concept 1 shows the basics of an electric dipole. The equal and opposite charges have magnitude q and are separated by the distance d. An electric dipole can be described by its electric dipole moment $\mathbf{p} = q\mathbf{d}$, where $\mathbf{d}$ is the displacement vector from the negative charge to the positive charge. The dipole moment vector points in the same direction as the displacement vector.

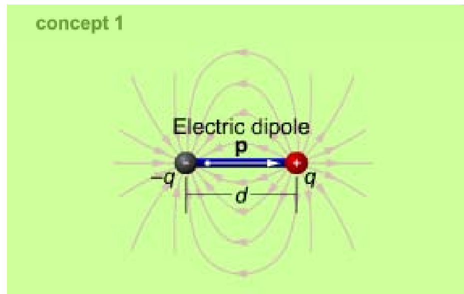
When an electric dipole is placed in an external electric field, the positive end experiences a force in the direction of the external field, and the negative end experiences a force in the opposite direction. As the diagram in Concept 2 shows, if the dipole moment is not aligned with the field, the dipole will experience a torque that rotates it into the direction of alignment with the external electric field.

The world's oceans contain about 7.8x10^{46} electric dipoles: water molecules.

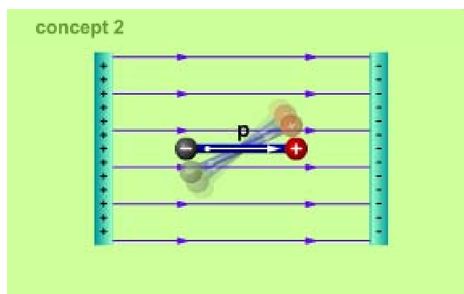
One reason why electric dipoles are important is that the individual molecules in what is arguably the most vital substance for human existence − water − behave as electric dipoles. The oxygen side of a water molecule tends to be negatively charged, while the hydrogen side tends to have an equal but opposite positive charge. The result is a dipole. The fact that water molecules are dipoles accounts for many of water's crucial properties. The illustration below shows a conceptual model of a water molecule together with its electric dipole moment.



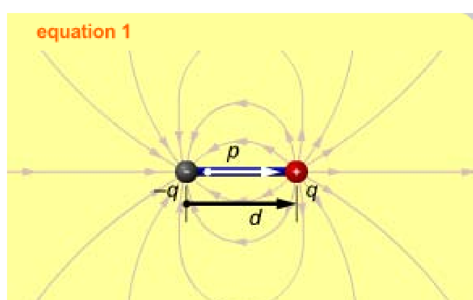
A convenient device, the microwave oven, takes advantage of the dipole nature of water molecules. The oven uses an oscillating electric field to cause water molecules to rotate back and forth, increasing their energy and the thermal energy of the food that contains them.



Electric dipole
Two point charges
· Separated by fixed distance
· Equal strength, opposite sign
Dipole moment vector p
· points from "–" to "+" charge



Electric dipole in an external field
Moment tends to align with external field



Dipole moment and electric field

$$\mathbf{p} = q\mathbf{d}$$

$\mathbf{p}$ = dipole moment

q = charge
d = displacement, "–" to "+" charge

24.18 - Gotchas

I increased the amount of the charge of a test charge, and the force exerted on it changed. This means the field I was assessing must have changed. No. Fields are independent of the test charge. The force increased, in proportion to the test charge, but the original field strength did not change. The increased force was a result of the increased magnitude of the test charge.

Electric field lines are the same as electric field vectors. No. The direction of the field is tangent to the field lines at any point. In a field diagram for a point charge, the lines are straight so they do point in the direction of the field. But in many fields the field lines are curved, and the field vector at any location on a field line is tangent to but distinct from it. Also, field lines and vectors use different conventions for representing field strength. With vectors, the length is proportional to the strength of the field. In a field diagram, the field lines are closer together in a region with greater field strength. Although a field diagram is used to represent an electric field, field lines and field vectors are not the same thing.

24.19 - Summary

An electric field is a vector quantity defined throughout a region of three-dimensional space. It describes the forces that will be experienced by an electric charge if it is placed at various locations in the field. If a positive charge is introduced, the force on it equals the field strength times its own charge magnitude, in the direction of the field. When a negative charge is introduced, the force on it is opposite to the direction of the field. The field strength is measured in newtons per coulomb (N/C). A test charge is a small positive charge used to measure a field.

Point charges generate electric fields. At any distance from the charge, the strength of the field is proportional to the magnitude of the charge and inversely proportional to the square of the distance from the charge.

An electric field diagram is a convenient means of representing the direction and strength of an electric field in a region. The direction of a field line represents the local direction of the electric field. The strength of the field at a particular location is indicated by the proximity of the field lines to each other around that location.

When more than one charge is present, electric fields obey the principle of superposition. This means electric fields at any point can be added as vectors.

In an isolated, charged conductor the excess charges distribute themselves on the conductor's surface in a state of electrostatic equilibrium. There is no net motion of these charges. The electric field extends perpendicularly outward from all points on the conductor's surface, and there is no field inside the conductor.

An electric dipole consists of a positive and a negative charge of equal magnitude, separated by a fixed distance. A vector called the dipole moment **p** points from the negative to the positive charge. Its magnitude is equal to the magnitude of either dipole charge times the distance between the charges. When the dipole is placed in an external electric field, the dipole moment experiences a torque tending to align it in a direction parallel to the field.

Dipoles generate electric fields. At large distances from the dipole, the strength of the field is proportional to the magnitude of the dipole moment and inversely proportional to the cube of the distance from the dipole. The equations on the right apply to a dipole centered at the origin, and oriented toward the right, along the x axis. This is the dipole axis; the y axis is the bisector axis.

Equations

Electric field

$E = F/q_{\text{test}}$

Field of a point charge

$$E = \frac{k|q|}{r^2}$$

Force due to an electric field

$\mathbf{F} = q\mathbf{E}$

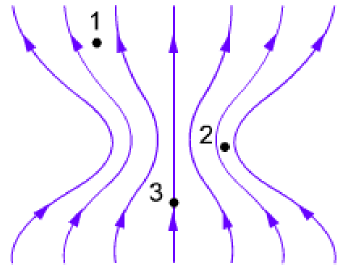
Electric dipole

$\mathbf{p} = q\mathbf{d}$

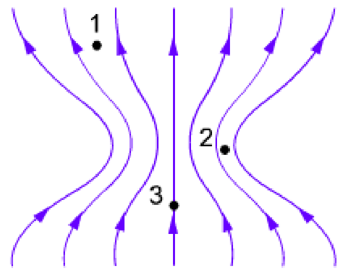
Conceptual Problems

C.1 Three locations, 1, 2, and 3, are labeled in the electric field diagram. At which of these locations is the strength of the field the greatest?

i. 1
ii. 2
iii. 3
iv. All three same strength

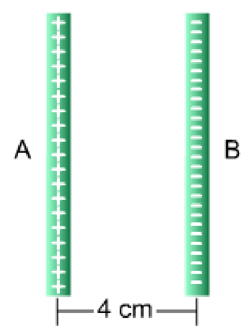


C.2 Suppose a charged particle is placed in the field whose diagram is shown here. (a) If a positively charged particle is placed at point 1, in what direction will it accelerate? (b) If a negatively charged particle is placed at point 3, in what direction will it accelerate? (c) At which of the three locations labeled in the diagram should a charged particle be placed so that it experiences the greatest acceleration?



(a) i. In the direction of the field
ii. Opposite the direction of the field
(b) i. In the direction of the field
ii. Opposite the direction of the field
(c) i. 1
ii. 2
iii. 3

C.3 The electric field created by a uniformly charged infinite plane is uniform in strength and perpendicular to the plane. Charged planes A and B are identical, except that A has a positive charge and B has the same magnitude of negative charge. Plane A is 4.0 cm to the left of plane B, and the two planes are parallel. (a) What direction is the combined field to the left of A? (b) Halfway between A and B? (c) To the right of B?



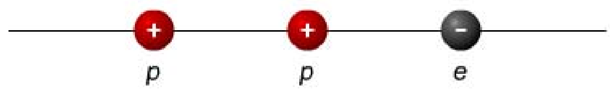
(a) i. To the right
ii. To the left
iii. The combined field is zero
(b) i. To the right
ii. To the left
iii. The combined field is zero
(c) i. To the right
ii. To the left
iii. The combined field is zero

C.4 A regular tetrahedron is a three-sided pyramid whose base and three sides are all equilateral triangles. So, a tetrahedron has four faces and four vertices, each vertex being directly opposite a face. If identical charges of $+q$ are placed at three of the vertices, what charge must be placed at the fourth vertex to ensure that the strength of the electric field at the center of the tetrahedron is zero?

○ 0 ○ q ○ $3q$ ○ $-q$ ○ $-3q$

C.5 In the diagram, two protons and one electron are placed on a line. Describe the part(s) of the line, relative to the positions of the particles, where the strength of the electric field due to the three charges can be zero. Select all that apply.

- ☐ Left of the protons
- ☐ Between the protons
- ☐ Between the rightmost proton and the electron
- ☐ Right of the electron



C.6 In the diagram, two charges of $+2q$ and $-5q$ are placed on a line. (a) There is a point on the line where the strength of the electric field due to the two charges is zero. Describe where the point is, relative to the positions of the two charges. (b) Is there any point **not** on the line, where the strength of the electric field is zero?

(a) i. To the left of $+2q$
ii. At $+2q$
iii. Between $+2q$ and $-5q$
iv. At $-5q$
v. To the right of $-5q$

(b) Submit answer on paper

C.7 A neutron is a particle with the same mass as a proton, but no electric charge. An alpha particle consists of two protons plus two neutrons. A proton and an alpha particle decide to have a "drag race" starting from zero velocity in a certain uniform electric field **E**. Who wins? Why?

○ Proton ○ Alpha particle

C.8 A conducting spherical shell contains a smaller concentric solid conducting sphere. The inner sphere carries a negative charge, and the outer sphere carries a positive charge of equal total magnitude. (a) Is there an electric field inside the inner sphere? (b) Is there an electric field outside the outer sphere? (c) Describe the nature of the electric field between the two spheres.

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c) Submit answer on paper

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following question. If the you keep the particle's charge constant, does the field strength increase linearly as you move closer to the charge, or does it greatly increase at points near the charge?

○ Increases linearly
○ Greatly increases

0.2 Use the simulation in the interactive problem in this section to answer the following question. (a) If you increase the magnitude of the source charge, does the field strength at a given point increase, stay the same, or decrease? (b) In what direction does the field point when the source charge is positive? (c) At a given point, does the field point in the same direction for a negative source charge as it does for a positive source charge?

(a) i. Increase
ii. Stay the same
iii. Decrease

(b) i. Towards the source charge
ii. Away from the source charge

(c) ○ Yes ○ No

Section 1 - Electric fields

1.1 An electric field has a value of 350 N/C at a particular point in space, directed along the *x* axis. What is the force this field exerts at this point on a particle of charge (a) 5.00 μC? (b) −8.25 mC? (c) −525 C? (d) 5.34×10^{-15} C? Make sure to check the sign of your answer.

(a) ________ N
(b) ________ N
(c) ________ N
(d) ________ N

1.2 A test charge of 2.50 μC is placed at various points in an electric field where it experiences forces in the positive direction of 25.0 μN at point A, 250 μN at point B, and 2.50 mN at point C. (a) What is the electric field strength at those points? (b) What forces would the field exert on a particle with charge −2.50 μC at the same points? (c) What forces would the field exert on an electrically neutral particle at the same points?

(a) A ________ N/C
B ________ N/C
C ________ N/C
(b) A ________ N
B ________ N
C ________ N
(c) A ________ N
B ________ N
C ________ N

1.3 The electric field in the copper wire of a household circuit is oriented along the wire, so that it pushes electrons through the circuit. (In alternating current, the direction of this electric field reverses 120 times a second.) A typical value for the field strength is 0.00850 N/C. The charge of a mobile electron in the wire is -1.60×10^{-19} C. What is the magnitude of the force that the field exerts on the electron?

________ N

1.4 A hard rubber comb can be given a negative electric charge by rubbing it against a variety of materials, including human hair. Suppose an atmospheric ion having a charge of $+1.60\times10^{-19}$ C is resting at a point near the comb where the strength of the field is 973 N/C. (a) What is the magnitude of the force that the field exerts on the ion? (b) Will the ion move toward the comb or away from the comb?

(a) ________ N
(b) i. Toward the comb
ii. Away from the comb

1.5 The surface of the drum of a laser printer is negatively charged, except in certain "print black" regions, such as letterforms, that have been discharged by a laser beam. Negatively charged ink particles carrying a charge of 5.0 mC cling to the "print black" areas but are repelled by an electric field elsewhere on the drum whose strength at the surface equals 120,000 N/C. What is the magnitude of the repelling force felt by these particles?

________ N

Section 2 - Electric fields and Coulomb's law

2.1 Two friends are playing a version of proton golf where the hole is marked by a single proton. The first friend reads his meter, and declares he has a field strength of 23.5 N/C. The second friend looks at her meter and realizes she is three times as far away. What field strength does the second friend's meter read?

________ N/C

2.2 An ion harvester is scouring deep space for isolated xenon (Xe^+) ions. A xenon ion is singly ionized, and has a charge of +*e*. If the field meter reads 2.93e−7 N/C, how far away is it from the ion?

________ m

2.3 In a hydrogen atom in its lowest energy state, the single electron and the single proton are separated by a tiny distance called the Bohr radius, 5.29×10^{-11} m. The proton carries a charge $e = 1.60\times10^{-19}$ C. (a) What is the strength of the electric field generated by the proton at the Bohr-radius distance? (b) In what direction does the field point?

(a) ________ N/C
(b) i. Towards the proton
ii. Away from the proton

2.4 A solid conducting sphere with a radius of 35.4 cm contains a total charge of 5.46 mC, evenly distributed over its surface. (a) What is the direction of the electric field at its surface? (b) What is the strength of the electric field at its surface? Hint: use the shell theorem, which states that when calculating an electrostatic force or field outside a charged sphere, the sphere can be treated as though all of its charge resides at its center.

(a) i. Tangent to the surface
ii. Radially away from the center
iii. Radially toward the center

(b) ________ N/C

2.5 A solid nonconducting sphere has a radius of 50.0 cm. It has a uniform charge density of 3.00×10^{-3} C/m^3. What is the strength of the electric field (a) at a distance of 2.00 m from the center of the sphere? (b) at the surface of the sphere? (c) at a distance of 25.0 cm from the center of the sphere? (Hint: use the shell theorem, which states that when calculating an electrostatic force or field outside a sphere with a uniform charge, the sphere can be treated as though all of its charge resides at the center.)

(a) ________ N/C
(b) ________ N/C
(c) ________ N/C

Section 5 - Interactive problem: fields and forces

5.1 Use the simulation in the first interactive problem in this section to answer the following questions. (a) Is the force on the test charge greater where the field lines are closer together, or where they are farther apart? (b) What is the relationship between the direction of the force and the field lines?

(a) i. Closer together
ii. Farther apart
iii. The force is the same all over

(b) i. The direction of the force is perpendicular to the field lines
ii. The direction of the force is tangent to the field lines

5.2 Use the simulation in the second interactive problem in this section to answer the following questions. (a) At a particular location, how does the direction of the force exerted by the field on a positive charge differ from the force direction on a negative charge? (b) How many times larger is the force on the $+2q$ charge than on the $+q$ charge?

(a) i. Directions are the same
ii. Directions are opposite
iii. Directions are perpendicular to each other

(b) i. They are equal
ii. Twice as large
iii. Four times as large
iv. Half as large

Section 6 - Electric fields caused by multiple charges

6.1 Two charges, $q_1 = -3.00$ mC and $q_2 = 4.50$ mC, are located at −2.00 cm and 3.00 cm on the x axis, respectively. What are the magnitude and direction of the resulting field at (a) $x = -3.00$ cm? (b) $x = 0$ cm? (c) $x = 5.00$ cm?

(a) ________ N/C, i. to the right
ii. to the left

(b) ________ N/C, i. to the right
ii. to the left

(c) ________ N/C, i. to the right
ii. to the left

6.2 A particle with charge +7.88 μC is placed at the fixed position $x = 3.00$ m in an electric field of uniform strength 300 N/C, directed in the positive x direction. Find the position on the x axis where the electric field strength of the resulting configuration is zero.

________ m

6.3 A particle with a charge of + 4.95e−3 C is placed on the y axis at $y = 3.50$ cm. Calculate the charge on a second particle, placed at $y = -2.00$ m, that will result in a field whose strength at the origin is zero.

________ C

Section 7 - Sample problem: calculate the net field of two charges

7.1 Two particles are situated on the x axis. The particle q_1, with 34 excess electrons, is situated at the point x = 100 μm. The other particle q_2, with 17 excess electrons, is located at the origin. Give the x value of a point between the particles where the strength of the field they generate is zero.

__________ m

7.2 Positive charges with magnitudes 3.50 mC and 6.75 mC are located at the origin and at the point (10.00, 0) meters on the x axis, respectively. What are the strength and direction of the field they generate at the point (0, 10.00) meters on the y axis?

__________ N/C, __________ °

7.3 Charges of −4.05 mC and +3.85 mC are located on the x axis at the points (3.00, 0) m and (10.00, 0) m, respectively. What are the strength and direction of the field they generate at the point (2.50, 7.00) m? (Specify the direction as an angle between 0° and 360°.)

__________ N/C, __________ °

7.4 An arrangement of four point charges gives a field strength of 201 N/C in the +y direction at point P = (4, 1) m. If another charge of +7.80 μC is added at (−2, 5) m, what is the new field strength at point P?

__________ N/C

7.5 The electric field $\mathbf{E}_1$ is directed parallel to the x axis at all points in three-dimensional space, and its value at any point, in N/C, is equal to the z coordinate of the point. The vector field $\mathbf{E}_2$ is also directed parallel to the x axis at all points in three-dimensional space, but its value at any point is equal to z − 1. Describe all the points in three-dimensional space where these two fields combine to make a net field with zero strength.

7.6 Two electric fields in space are defined by $\mathbf{E}_1 = (3.0y, 5.4, 1.0x^2)$ N/C and $\mathbf{E}_2 = (2.3y, 2.5x, 0)$ N/C. Use the principle of superposition to compute the net electric field at the point (x, y, z) = (7.4, −3.2, 8.8). (a) What is the electric field at this point? (b) What is the net electric field at a general point (x, y, z)?

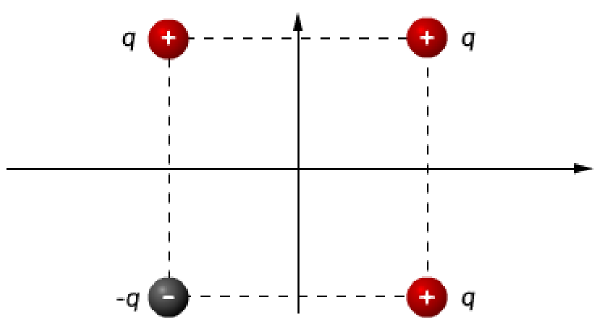
(a) (__________ , __________ , __________) N/C

(b) Submit answer on paper

Section 8 - Sample problem: calculate the net field of three charges

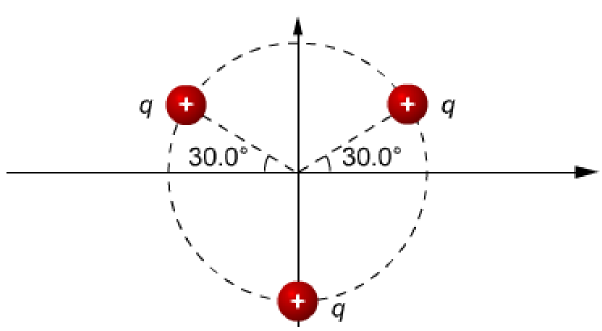
8.1 Four charges are placed at the vertices of a square, centered at the origin, as shown in the diagram. If each side of the square has a length of 0.224 m, what is the strength and direction of the net electric field at the origin? Express your answer in terms of the charge magnitude q.

__________ q N/C, __________ °

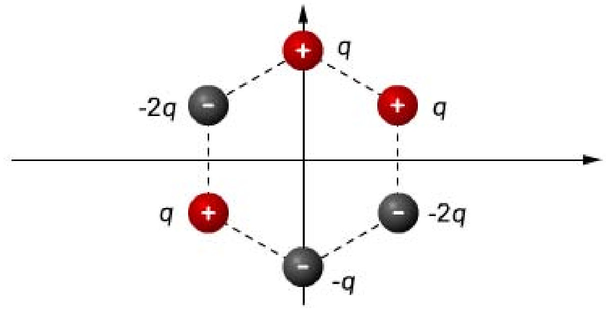


8.2 Three particles having the same charge q = 5.00 μC are arranged on a circle of radius 2.00 m, centered at the origin at the angles shown in the diagram. What is the magnitude of the net electric field at the origin due to these three charges?

__________ N/C



8.3 Six charges are at the vertices of a regular hexagon, centered at the origin, as shown in the diagram. If the magnitude of the charge q is 2.25 mC, and the distance of each charge from the origin is 5.00 cm, what are the magnitude and direction (as an angle between 0° and 360°) of the electric field at the origin due to these six charges?



__________ N/C, __________ °

8.4 Three charges of 3.02 μC, 7.36 μC, and 8.11 μC are located at the points (2.00, 7.00) m, (4.00, 7.00) m and (4.00, 3.00) m, respectively. (a) What is the magnitude of the combined field due to these charges at the point (2.00, 3.00) m? (b) What is the direction of the combined field there (as an angle between 0° and 360°)?

(a) __________ N/C
(b) __________ °

8.5 A particle with charge +3.46 mC is located at the point (0, 2.36) cm. Another particle with charge −5.32 mC is located at the point (4.65, 5.00) cm. Finally, a third particle with charge +2.00 mC is located at the point (−7.60, 0) cm. What are the components of the net field at the origin due to these three charged particles?

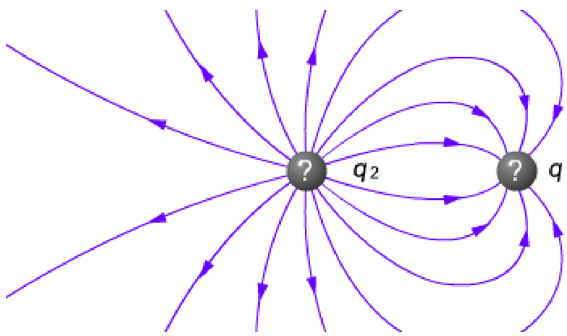
(__________ , __________) N/C

Section 9 - Drawing field diagrams for multiple charges

9.1 Two point charges of q and $-3q$ located near each other create an electric field. If a correctly drawn diagram of the field shows six field lines touching q, describe the electric field lines touching $-3q$.

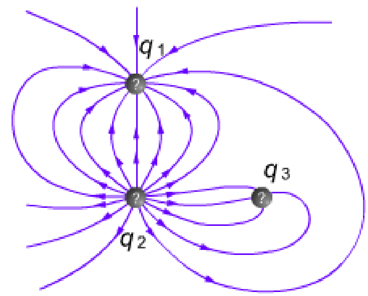
- ○ Two lines going out of $-3q$
- ○ Two lines going into $-3q$
- ○ 18 lines going out of $-3q$
- ○ 18 lines going into $-3q$

9.2 The illustration shows a diagram for the electric field created by two charged particles, q_1 and q_2. (a) Which charge is positive? (b) Which charge is negative? (c) Which charge has a greater magnitude? (d) How many times greater is it?



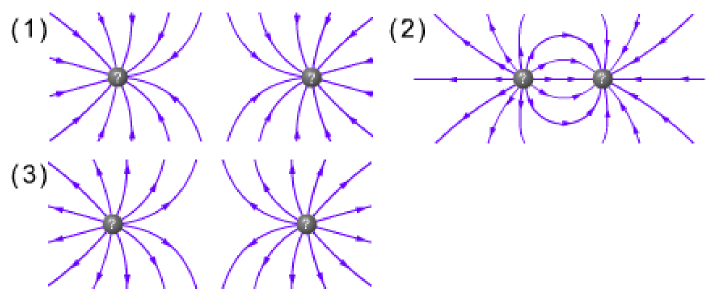
(a) ○ q_1 ○ q_2 ○ Neither ○ Both
(b) ○ q_1 ○ q_2 ○ Neither ○ Both
(c) ○ q_1 ○ q_2
(d) __________ times greater

9.3 The illustration shows a diagram for the electric field created by three charged particles q_1, q_2 and q_3. (a) Which charge(s) is/are positive? (b) Which charge(s) is/are negative? (c) List the charges in order of increasing magnitude (disregarding sign). (d) Charge q_1 is how many times stronger than charge q_3?



(a) ☐ q_1
☐ q_2
☐ q_3
(b) ☐ q_1
☐ q_2
☐ q_3
(c) ○ $|q_1| > |q_2| > |q_3|$
○ $|q_1| > |q_3| > |q_2|$
○ $|q_2| > |q_1| > |q_3|$
○ $|q_2| > |q_3| > |q_1|$
○ $|q_3| > |q_1| > |q_2|$
○ $|q_3| > |q_2| > |q_1|$
(d) __________ times stronger

9.4 The three electric field diagrams labeled "1", "2", and "3" in the accompanying illustration each represent the field around two charged particles of equal magnitude. (a) Which diagram has a positive charge and a negative charge? (b) Which diagram has two positive charges? (c) Which diagram has two negative charges?



(a) ○ 1 ○ 2 ○ 3
(b) ○ 1 ○ 2 ○ 3
(c) ○ 1 ○ 2 ○ 3

9.5 Two point charges, $+4q$ and $-q$, are located near each other. Sketch a diagram of the electric field generated by these two charges, including some indication of how the field lines behave between the particles as well as at greater distances from them.

9.6 Two point charges, $+4q$ and $+q$, are located near each other. Sketch a diagram of the electric field generated by these two charges, including some indication of how the field lines behave between the particles as well as at greater distances from them.

Section 10 - Describing the force exerted by an external electric field

10.1 A particle of mass 1.20 kg with a charge of +2.30 mC is shot into a region with an electric field of strength 18.0 N/C, having a direction opposite the particle's velocity. If the particle has an initial speed of 3.10 m/s, how long does it take to come to a complete (though momentary) standstill? Assume there are no other forces acting on the particle.

__________ s

10.2 Protons carry an electric charge of $e = 1.60 \times 10^{-19}$ C, electrons carry a charge of $-e$, and neutrons are electrically neutral. Protons and neutrons have a mass of 1.67×10^{-27} kg, and electrons have a mass of 9.11×10^{-31} kg. A uniform electric field can exert a force on a charged particle that effectively neutralizes the force of gravity on the particle, causing it to be "weightless." What is the strength of the electric field that will levitate (a) a proton? (b) an electron? (c) an alpha particle (two protons and two neutrons)?

(a) __________ N/C
(b) __________ N/C
(c) __________ N/C

10.3 An electron (charge -1.60×10^{-19} C, mass 9.11×10^{-31} kg) is in a uniform electric field of strength 543 N/C in the positive direction. (a) What is the force on the electron? (b) What is the acceleration of the electron? (c) What is the electron's speed after 1.05×10^{-8} s? (d) What fraction of the speed of light (3.00×10^{8} m/s) is your answer to part c? Be careful of the sign in each of your answers.

(a) ________ N
(b) ________ m/s^2
(c) ________ m/s
(d) ________

10.4 A proton (charge 1.60×10^{-19} C, mass 1.67×10^{-27} kg) is in a uniform electric field of strength 2.51×10^{4} N/C. (a) What is the magnitude of the force on the proton? (b) What is the magnitude of acceleration of the proton? (c) What is the proton's speed after 1.25 ns?

(a) ________ N
(b) ________ m/s^2
(c) ________ m/s

10.5 An alpha particle (charge 3.20×10^{-19} C, mass 6.68×10^{-27} kg) is in a uniform electric field of strength 7.75×10^{-3} N/C. What is the particle's speed after 12.5 ms?

________ m/s

10.6 A proton with mass 1.67×10^{-27} kg and charge 1.60×10^{-19} C is traveling at 6.00×10^{5} m/s in the positive direction when it enters a uniform electric field with a strength of 1250 N/C in the negative direction. The opposing electric force brings the proton to rest. Calculate the displacement of the proton while it is coming to rest.

________ m

10.7 A proton with mass 1.67×10^{-27} kg and charge 1.60×10^{-19} C accelerates from rest in a uniform electric field of strength 500 N/C. (a) What is the magnitude of the acceleration of the proton? (b) How long does it take the proton to reach a speed of 35,000 m/s? (c) What distance has the proton traveled when it reaches this speed? (d) What is the kinetic energy of the proton at 35,000 m/s?

(a) ________ m/s^2
(b) ________ s
(c) ________ m
(d) ________ J

Section 12 - Interactive problem: the alpha cannon

12.1 In the description of the alpha cannon interactive problem in the textbook, we claim that the force exerted on an alpha particle by the uniform electric field in the simulation is so much greater than the force of gravity on it that gravity can be safely ignored in solving the problem. Here you verify this claim. How many times larger is the electric force than the force of gravity?

________ times larger

12.2 Using the information given in the interactive problem in this section, what initial velocity is required for the alpha particle to hit the aluminum atom? Test your answer using the simulation.

________ m/s

12.3 Using the electric field described in the alpha cannon simulation, suppose that an alpha particle is fired upward from a height y = 0 m, with horizontal and vertical velocity components v_x = 15.0 m/s and v_y = 5.00 m/s. (a) How much time passes before the particle returns back down to the height y = 0 m? (b) What is the horizontal distance traveled by the particle before it returns to the height y = 0 m?

(a) ________ s
(b) ________ m

12.4 Using the electric field described in the alpha-cannon simulation, suppose that an alpha particle is fired upward and forward from a height y = 0 m at a speed v = 28.0 m/s. The target is an aluminum atom to the right of the alpha particle, at the same height y = 0 m, and at a distance x = 4.00 cm. There are two possible firing angles that will result in a hit: an arcing "lob" shot at a high angle; and a more direct "forward" shot at a lower angle. (a) At what angle above the horizontal should the alpha particle be fired to hit the aluminum atom with a lob shot? (b) At what angle above the horizontal should the alpha particle be fired to hit the aluminum atom with a forward shot?

(a) ________ °
(b) ________ °

Section 14 - Sample problem: field strength of an ion drive

14.1 A particle with a mass of 3.65×10^{-24} kg and a charge of 6.40×10^{-19} C starts at rest and is accelerated for a distance of 0.0960 m through a uniform electric field having strength $E = 7.25\times10^4$ N/C. What is the resulting speed of the particle?

_________ m/s

14.2 An astronaut performs a simple experiment in the "weightless" environment of a space station, intended for television broadcast to elementary school classrooms around the world. By combing her hair, she is able to charge a flat comb so that, in the vicinity of the comb, there is an approximately uniform electric field of strength $E = 900$ N/C. She places the comb 4.00 cm away from a clump of small paper dots, inducing a nonuniform distribution of charge in the clump so that one of the dots becomes charged and breaks away, accelerating toward the comb. If the paper dot has a mass of 2.95×10^{-4} kg and a charge of magnitude 3.50 μC, how fast is it going when it strikes the comb?

_________ m/s

Section 15 - Interactive problem: tune the rocket’s drive field

15.1 Using the information given in the interactive problem in this section, what is the uniform electric field strength necessary to accelerate the ions, from rest, to the desired exhaust velocity? Test your answer using the simulation.

_________ N/C

Section 17 - Electric dipoles

17.1 Two opposite point charges of magnitude 2.30e−6 C are located near the origin. The positive charge is at (1.50, 1.50) mm and the negative charge is at (−1.50, −1.50) mm. What is the magnitude of the dipole moment vector for this pair of charges?

_________ C · m

17.2 The magnitude of the dipole moment of a water molecule in liquid water is 6.2×10^{-30} C·m. A simplified model of the water molecule H_2O can be constructed by imagining that each of the two hydrogen atoms in the molecule contributes its single electron to the one oxygen atom, and thereby becomes bound to it. In this model, the molecule is a simple dipole, with the oxygen atom at one end and the two hydrogen atoms at the other, each end having a charge of magnitude $2e = 3.2\times10^{-19}$ C. In this model, what is the dipole separation of the charges in the water molecule?

_________ m

17.3 An electric dipole with moment **p** = (3.43, −5.63) μC·m consists of two opposite charges of magnitude $1.00\text{x}10^{-6}$ C. How far apart are the charges?

_________ m

Additional Problems

A.1 When a test charge of 6.0×10^{-10} C is placed at the origin, it experiences a force of 2.8e−3 N in the negative *x* direction from an external electric field. (a) What is the magnitude of the electric field at the origin? (b) If this field is due to a point charge at $x = 4.0$ m on the *x* axis, what is the amount of this charge?

(a) _________ N/C
(b) _________ C

A.2 A charge of 6.00 μC is placed at the origin. (a) Find the electric field at the point (3.00, 4.00) m. (b) What force does a test charge of 3.50 μC experience at the point (10.0, −3.00) m? Express your answers as vectors in polar notation with an angle between 0° and 360°.

(a) (_________ N/C, _________ °)
(b) (_________ N, _________ °)

A.3 There is an electric field of about 130 N/C near the surface of the Earth during fair weather. This electric field is directed towards the surface of the Earth. (a) Is the excess charge on the surface of the Earth positive or negative? (b) If a plastic foam sphere has -3.50×10^{-5} C of charge on it, what is the magnitude and direction of the force on the sphere from the Earth's electric field?

(a) ○ Positive ○ Negative
(b) _________ N
i. Toward the Earth
ii. Away from the Earth

A.4 A physicist wants to use an electric field to accelerate an electron to a speed of 1.00% of *c*, where *c* is the speed of light. If his accelerator can produce a uniform electric field of 5.00×10^6 N/C, how long does the accelerator tube have to be? (The speed of light is 3.00×10^8 m/s, the charge of an electron has magnitude 1.60×10^{-19} C and the mass of an electron is 9.11×10^{-31} kg. Ignore relativistic effects.)

_________ m

A.5 Van de Graaff accelerators use static electricity to accelerate particles to high energies. A typical electric field in a Van de Graaff accelerator is 2.50×10^{6} N/C. (a) What is the magnitude of the force on an electron in this field? (b) How about on an alpha particle, which carries a charge of $2e$?

(a) __________ N
(b) __________ N

A.6 A point charge sits in an otherwise charge-free region. Points 1 and 2 are 3.00 cm and 6.00 cm from the charge, respectively. If the electric field at point 1 is 3.60e−5 N/C, what is the magnitude of the electric field at point 2?

__________ N/C

A.7 A charge of −3.50 μC is located on the *x* axis at *x* = −4.00 m, and a second point charge of 4.00 μC is located at *x* = 6.00 m. Where should a third point charge of 6.00 μC be placed to make the electric field zero at the origin?

__________ m

A.8 When a "pith ball" (a small, lightweight nonconducting ball of mass m capable of holding a relatively large static charge) is suspended with a thin thread of length L, it will swing as a pendulum whose period is given by

$$T = 2\pi\sqrt{\frac{L}{g}} = 2\pi\sqrt{\frac{L}{F_g/m}} = 2\pi\sqrt{\frac{Lm}{F_g}},$$

where F_g is the force of gravity on the ball. Now suppose that the thread has a length of 7.50 cm, and the ball has a mass of 8.00 grams and a charge of 3.10 μC. (a) If the pendulum is placed in a uniform electric field of 1.26×10^{4} N/C, directed upward, what will its period be? (b) What should the strength of the field be to slow the pendulum's period to 2.13 seconds?

(a) __________ s
(b) __________ N/C

chapter 25 Electric Potential

25.0 - Introduction

In this chapter, we discuss electric potential energy and electric potential. Understanding these topics will be crucial to your understanding of how electric circuits and components work. For instance, when you turn on a flashlight, you are allowing an electric potential difference to drive a current that causes the light bulb to glow.

In the sections that follow, you will learn how electric charges and fields can create electric potential energy, electric potential, and electric potential differences. You may be unfamiliar with the term electric potential difference, but you likely know its units, volts, and you have probably heard the informal term “voltage” used for it.

The simulations to the right allow you to experiment with electric potential energy. Just as a configuration of masses, such as a barbell held above the Earth’s surface, possesses gravitational potential energy, so too does a configuration of electrically charged particles possess electric potential energy.

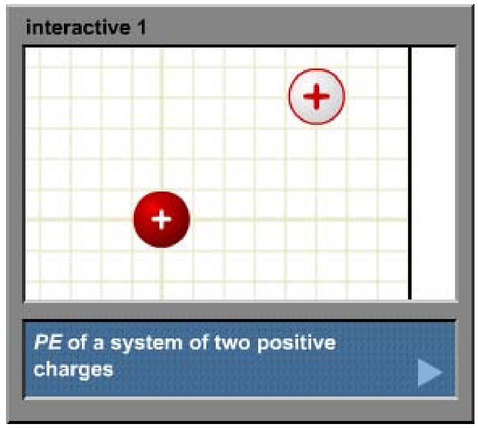
The first simulation contains a stationary positive charge (red) and a positive test charge (white) that you drag around with your mouse. As you move the test charge from place to place, you can change the electric potential energy of the system of two charges. The potential energy, which depends on the distance r between the charges and their strengths, is displayed in the control panel. A graph of the PE is drawn on the right side of the simulation as you move the test charge around. A grid in the background of the simulation allows you to estimate the distance in millimeters between the two charges.

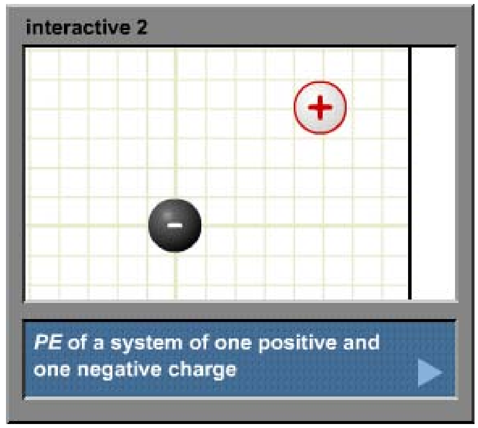
As you move the test charge, consider the following questions: What is the sign of the PE? When is the potential energy the greatest? The least? The graph shows PE as a function of the distance r between the charges.

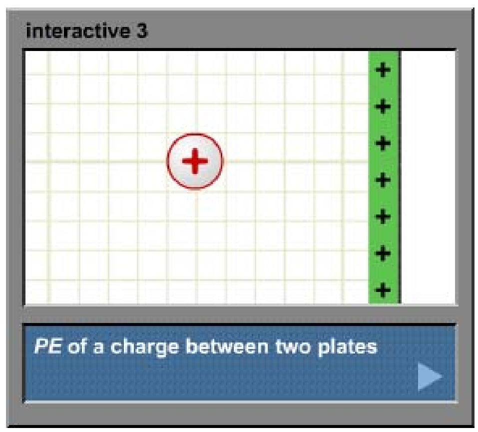
As with gravitational potential energy, a configuration having zero potential energy must be defined. In the first two interactives, the potential energy is defined to be zero when an infinite distance separates the two charges.

Interactive 2 is the same as Interactive 1 except that the stationary charge is negative instead of positive. Experiment with this configuration by moving the positive test charge around. What is the sign of the PE now? As the distance increases, does the PE increase or decrease? Can you move the charge in a way such that the PE does not change?

In the Interactive 3 simulation, you see a test charge that you can move between two oppositely charged plates. The electric field is uniform between the plates. In this simulation, the x axis of the graph tells the distance of the test charge from the negative plate. How does the PE change as you move the test charge away from the negative plate and toward the positive plate?







25.1 - Electric potential energy

Electric potential energy: **Potential energy determined by the configuration of electric charges.**

Electric potential energy, PE_e, is the measure of the energy stored due to the configuration of a system of charges. It reflects the positions (not the motion) of the charges in the system.

The release of electric potential energy can be highly visible. In a thunderstorm, clouds often accumulate a tremendous number of charged particles in a configuration that stores huge amounts of electric potential energy. Part of this energy can be released in the form of a lightning strike, a particularly dramatic example of the effects of discharging PE_e.

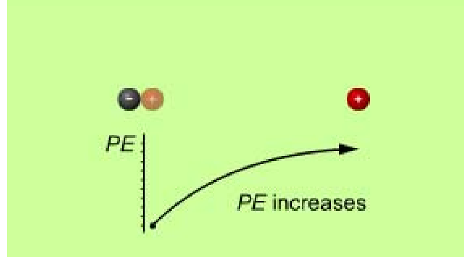
Recalling the fundamentals of gravitational potential energy may help you to understand electric potential energy because the two are analogous. For example, the distance of an apple above the Earth’s surface directly affects its gravitational potential energy, or more precisely, the gravitational potential energy of the Earth/apple system. Lifting the apple higher requires work − a force applied through a distance − and increases the gravitational potential energy. Lowering the apple reduces the potential energy of the system.

This Honda FCX automobile can cruise up to 350 km by using the electric potential energy stored in its capacitors as a configuration of electric charge.

To apply these concepts to electric potential energy, consider the initial and final configurations of opposite charges shown in the illustration of Concept 1. Initially, the charges are touching. Then, the positive charge is moved to the right by an external force while the negative charge on the left is held stationary. The work done on the system by the force increases its potential energy.

In this chapter, since we are discussing potential energy in the context of electro**statics**, we focus on stationary charges. The charges are stationary in their initial and final configurations, which means their KE does not change. Any work done on the charges contributes only to changing their PE_e. We state that as the first equation in Equation 1: Work done on the charges equals the change of PE_e. You can think of a direct gravitational analogy: The work you do to raise an apple from one stationary position to another equals its increase in PE.

concept 1

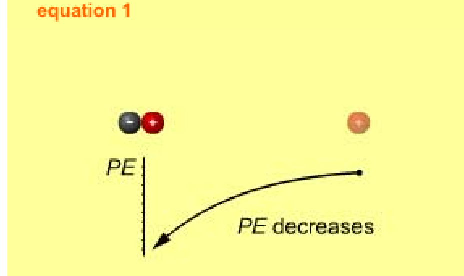


Electric potential energy
Property of a system of charges
Depends on charge separation, strength

Two charges with opposite signs attract each other, just as two masses do. Separate two such charges, or two masses, and you increase their PE_e. However, like charges repel each other. It takes work to move them closer together, and this means their PE_e increases as they approach each other. The simulations you used in the introduction to this chapter are designed to emphasize these points, and this may be a good time to try them again. They show the similarities and differences in the PE_e of similar configurations of like or opposite charges.

So far, we have been discussing work done **on** a system of charges by an external force. On the other hand, the fields of the particles themselves can do work on each other, similar to what happens when an apple falls to the Earth under the pull of gravity. This is referred to as work being done **by** the system. You see an example of this in the illustration of Equation 1: After positive work done on the system separates the charges and increases its potential energy, positive work done by the system pulls them back together and **decreases** its potential energy again. This is directly analogous to the relationship of work and gravitational potential energy: The gravitational potential energy of a system decreases when two objects approach, as when an apple falls to the Earth. We state this relationship between ΔPE_e and work as the second equation in Equation 1: The change in PE_e equals the opposite of the work done **by** the system.

equation 1



Changes in energy

$\Delta PE_e = W$ (work done **on** system)

PE_e = electric potential energy
W = work

$\Delta PE_e = -W$ (work done **by** system)

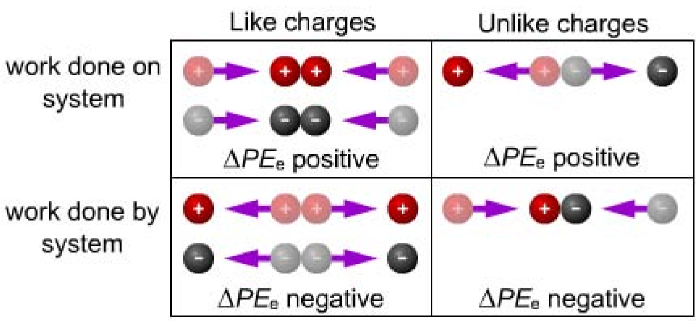
Charges are stationary at end points

In either case − work done on the system or by the system − the work equals a force applied through a displacement. The electrostatic force increases with the strength of the charges, and it decreases with distance. Be careful when you are asked to find the work done in changing the separation between two charges. The amount of force constantly varies with distance, a fact you must account for when you calculate the work. Calculus proves a useful tool for this, since it provides a technique for calculating work as the force varies over each small increment of displacement.

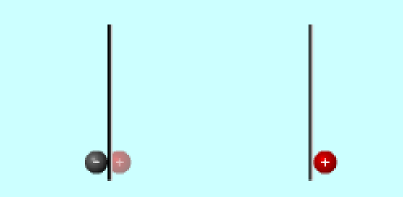
The electrostatic force is conservative, and the work done by an electric field on charges is path independent. This means that the work done by the field as a charged particle moves from one stationary position to another does not depend on the particle's path between the positions. For example, in the illustration of Concept 1, a horizontal force was exerted to push one particle directly away from the other to increase the system's potential energy. If the two particles had been moved from their indicated starting positions to their end positions by other paths (perhaps involving extravagant zigzags and loops), the net work done and the final potential energy would have been the same. The conservative nature of the electrostatic force means that the potential energy depends solely on the configuration, not on how it was arrived at.

We conclude this section with a review of the relationship of the signs of work and potential energy, since they can be confusing. In terms of work done **on** the system, **separating** two charges with **opposite** signs takes positive work (force in the direction of displacement) and increases their potential energy. **Pushing together** two **like** charges also takes positive work, and also increases their potential energy. Checking a computationally based answer against some of your own physics intuition may help make this clear. For instance, if you have to pull apart opposite charges, it should seem that you are doing work on the system, increasing its energy, and in fact you are.

The proper treatment of signs is summarized in the table below.



example 1



It takes 5.0 J of work by the right-hand wand to separate the charges. What is the change in potential energy?

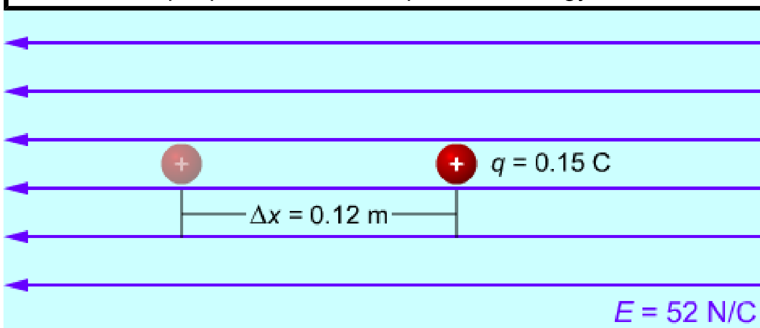
$\Delta KE = 0$

$\Delta PE_e = W$ (work done **on** system)

$W = 5.0\text{ J}$

$\Delta PE_e = 5.0\text{ J}$

25.2 - Sample problem: electric potential energy



A particle with a charge of 0.15 coulombs is in a uniform electric field of strength 52 N/C. An external force pushes the charge 0.12 meters directly against the field.

What is the change in electric potential energy?

In this problem, work is done **on** the system consisting of the charged particle q and the uniform electric field E. The particle is stationary both before and after it is moved.

Variables

change in electric potential energy	ΔPE_e
work done to move particle	W
change in particle's kinetic energy	$\Delta KE = 0\text{ J}$
force moving particle	F
charge of particle	$q = 0.15\text{ C}$
strength of electric field	$E = 52\text{ N/C}$
distance particle moves	$\Delta x = 0.12\text{ m}$

What is the strategy?

1. Calculate the work done in moving the particle from its initial to its final position.
2. Since there is no change in kinetic energy, the change in potential energy equals the work done on the system. Using the given values, calculate this change.

Physics principles and equations

The work done to move the particle against the field is

$$W = F\Delta x$$

The force required to move the particle against the field is

$$F = qE$$

Since the charged particle is stationary before and after it is moved by the external force, $\Delta KE = 0$. This allows us to apply the equation relating the change in potential energy to work done on the system,

$$\Delta PE_e = W$$

Step-by-step solution

We calculate the work done to move the charge q through the displacement Δx.

Step	Reason
1. $W = F\Delta x$	definition of work
2. $F = qE$	definition of field
3. $W = qE\Delta x$	substitute equation 2 into equation 1

Since $\Delta KE = 0$, we may state that ΔPE_e equals the work done on the system, and evaluate the result.

Step	Reason
4. $\Delta KE = 0$	particle stationary before and after
5. $\Delta PE_e = W$	work done on system
6. $\Delta PE_e = qE\Delta x$	substitute equation 3 into equation 5
7. $\Delta PE_e = (0.15\ \mathrm{C})(52\ \mathrm{N/C})(0.12\ \mathrm{m})$ $\Delta PE_e = 0.94\ \mathrm{J}$	evaluate

The solution shows that the system has a greater potential energy when q is in its final position. The external force did positive work on the system, increasing its PE.

25.3 - Electric potential energy and work

For a given configuration of a system to have a certain “absolute” amount of potential energy, a reference configuration with zero potential energy must be established. Changes from that zero point will then be the measure of the potential energy of the system.

With gravitational potential energy, an infinite separation between two masses is often defined as the configuration that has zero potential energy. This convention is frequently used for orbiting bodies. It is also common to use the convention that an object at the Earth’s surface represents a configuration with zero gravitational potential energy.

When analyzing electric potential energy, physicists usually state that a system of two charges has zero electric potential energy when an infinite distance separates them. The relationship between ΔPE_e and work that was introduced previously can be combined with this reference value to define a relationship between absolute PE_e and work, shown in Equation 1 on the right.

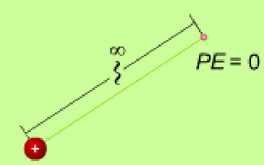
The relationship between the work W_∞ done **by** a system of two charges as one of them moves in from an infinite separation, and the PE_e of the system, is $PE_e = -W_\infty$. Why does this equation hold true? We already know that $\Delta PE_e = -W$, and when the initial PE_e is zero, then the change in potential energy is just the potential energy of the system in its final configuration.

In Equation 2, we show you how to calculate the work done by the system as the separation between charges changes. This equation can be derived using calculus. Note that the signs of the charges **do** matter in the formula. If they are opposite, applying the equation confirms that the system (the field) does positive work as the particles approach one another. Conversely, the work done by the field is negative if two like charges are moved closer to one another, as illustrated in the diagram.

In Equation 3, we combine the work and potential energy equations using an infinite separation to define zero potential energy. Since zero is the result of dividing by infinity, the $1/r_i$ term “disappears” and only the fraction $-1/r_f$ remains in the final factor of the work formula. Its negative sign cancels the one in the potential energy equation $PE_e = -W_\infty$, to yield the equation you see as Equation 3.

Note an implication of this equation: When the signs of the two charges are opposite, the potential energy of the pair is negative. When they are the same, the potential energy is positive.

concept 1



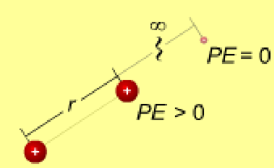
Electric potential energy

To determine electric potential energy

· Set point of zero potential energy

· Infinite separation often used

equation 1



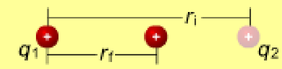
Electric potential energy and work

$$PE_e = -W_\infty$$

PE_e = electrical potential energy

W_∞ = work done **by** system as charge moves in from infinity

equation 2



Work by field

$$W = kq_1q_2\left(\frac{1}{r_i} - \frac{1}{r_f}\right)$$

W = work done **by** field
k = Coulomb's constant
q = electric charge
r_i = initial separation
r_f = final separation

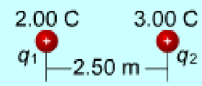
equation 3

Electric potential energy
When PE_e = 0 at infinite separation

$$PE_e = \frac{kq_1q_2}{r}$$

PE_e = electric potential energy
k = Coulomb's constant
q = electric charge
r = separation of charges

example 1



What is the potential energy of this system?

$$PE_e = \frac{kq_1q_2}{r}$$

$$PE_e = \frac{(8.99\times 10^9 \frac{\text{N}\cdot\text{m}^2}{\text{C}^2})(2.00\text{C})(3.00\text{C})}{2.50\text{ m}}$$

$$PE_e = 2.16\times 10^{10}\text{ J}$$

25.4 - Electric potential energy: multiple point charges

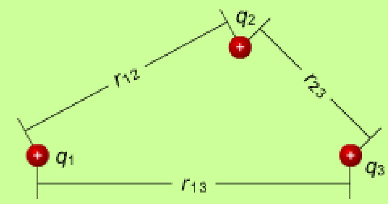
The electric potential energy of a system of three or more point charges can be determined, though it takes a bit more work than determining the potential energy of a two-charge system. In this section, we describe how to calculate the electric potential energy of a three-charge configuration, and state a formula for any number of charges.

To calculate the potential energy of a system of charges, start by selecting one charge. This charge has no electric potential energy by itself. Then calculate how much potential energy you generate by bringing a second charge near the first. For the third charge, calculate how much potential energy bringing it near the first charge generates, and also calculate how much potential energy bringing it toward the second charge creates. Add these three potential energies and you will have the overall potential energy of the configuration.

This summing of energies is allowed because of the principle of superposition. The forces applied by the system as you assemble this configuration can all be added, meaning that the work performed and the amounts of potential energy that result can be added, as well.

In Equation 1 you see formulas for the potential energy of a three-charge configuration, as well as a general formula for a configuration of n charges. Note the special summation symbol in the second formula. It indicates that you are supposed to add up fractions like the one shown in the sum for all values of i and j that have $i < j$. This prevents "double counting" the potential energy of each charge pair. You see a particular example of this rule, for three charges, in the first formula.

concept 1



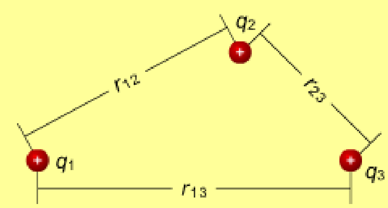
Electric potential energy: multiple point charges

Assemble system one charge at a time

· Determine *PE* contributed by each

· Sum potential energies of all pairs

equation 1



Total *PE* is sum of *PE*'s from each charge pair

For a system of 3 charges:

$$PE_e = \frac{kq_1q_2}{r_{12}} + \frac{kq_1q_3}{r_{13}} + \frac{kq_2q_3}{r_{23}}$$

PE_e = net electric potential energy

k = Coulomb's constant

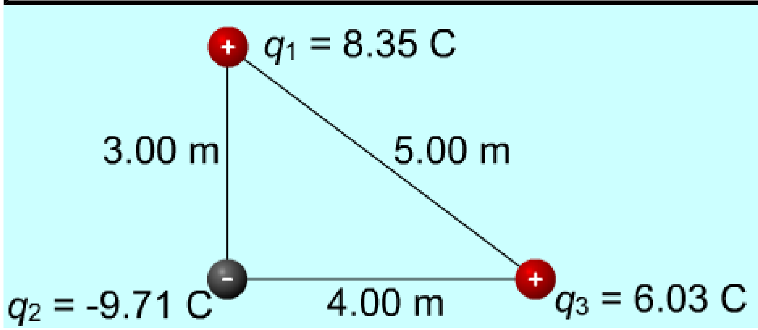
q = signed charge

r = separation of charges

For a system of n charges:

$$PE_e = \sum_{i<j} \frac{kq_iq_j}{r_{ij}} = k\sum_{i<j} \frac{q_iq_j}{r_{ij}}$$

25.5 - Sample problem: multiple point charges



What is the total electric potential energy (PE_e) of this configuration of three charges?

Above, you see three charges q_1, q_2 and q_3 located at the vertices of a triangle.

Variables

total electric PE of system	PE_e
a charge	q_i
electric PE of charge pair i, j	PE_{ij}
separation of charges i and j	r_{ij}
Coulomb's constant	$k = 8.99\times10^9\ \mathrm{N{\cdot}m^2/C^2}$

We will carry the constant k through our work until the final calculation without substituting its numeric value, in order to simplify the presentation of the steps below.

What is the strategy?

1. Start with no charges and zero PE_e in the configuration. Place the first charge in its position. The potential energy of just one charge is zero.
2. Add the second charge to the configuration and calculate the potential energy of the charge pair you have created.
3. Add the third charge, calculate the potential energies of the two new charge pairs you have created, and add them to the previous PE_e. The result will be the total potential energy of the illustrated configuration.

Physics principles and equations

The principle of superposition allows us to state that the total electric potential energy of a configuration of charges equals the sum of the potential energies for each pair of charges.

$$PE_e = PE_{12} + PE_{13} + PE_{23}$$

The potential energy of a pair of charges q_i and q_j is

$$PE_{ij} = \frac{kq_iq_j}{r}$$

Be careful with signs in this problem!

Step-by-step solution

One charge q_1 has no potential energy. When a second charge q_2 is joined to the configuration, the total potential energy is PE_{12}.

Step	Reason
1. q_1 only: $PE_e = 0$	single charge has no PE
2. add q_2: $PE_{12} = \dfrac{kq_1q_2}{r_{12}}$	equation for PE
3. $PE_{12} = \dfrac{k(8.35\ \mathrm{C})(-9.71\ \mathrm{C})}{3.00\ \mathrm{m}}$ $PE_{12} = -27.0k\ \mathrm{C^2/m}$	evaluate

When a third charge q_3 is joined to the configuration, the total potential energy is increased by two additional terms: PE_{13} and PE_{23}.

Step	Reason
4. add q_3: $PE_{13} = \frac{kq_1q_3}{r_{13}}$ $PE_{23} = \frac{kq_2q_3}{r_{23}}$	equations for PE
5. $PE_{13} = \frac{k(8.35\text{ C})(6.03\text{ C})}{5.00\text{ m}}$ $PE_{13} = 10.1k\text{ C}^2/\text{m}$	evaluate
6. $PE_{23} = \frac{k(-9.71\text{ C})(6.03\text{ C})}{4.00\text{ m}}$ $PE_{23} = -14.7k\text{ C}^2/\text{m}$	evaluate

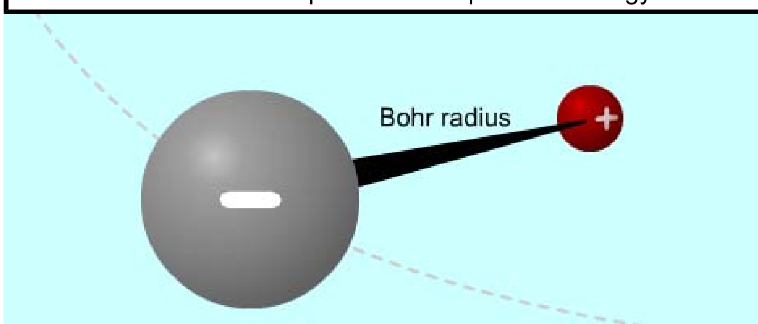
Now we add up the potential energy contributions calculated above for the three pairs of charged particles.

Step	Reason
7. $PE_e = PE_{12} + PE_{13} + PE_{23}$	superposition
8. $PE_e = -27.0k + 10.1k + -14.7k\text{ C}^2/\text{m}$ $PE_e = -31.6k\text{ C}^2/\text{m}$	substitute equations 3, 5, and 6 into equation 7
9. $PE_e = (-31.6\text{ C}^2/\text{m})(8.99\times10^9\text{ N}\cdot\text{m}^2/\text{C}^2)$ $PE_e = -2.84\times10^{11}\text{ J}$	enter value for k and multiply

The procedure followed above, which starts with a single PE term and then adds two more, does not have to stop with just three charges. Imagine adding a fourth charge q_4 to the configuration. The total potential energy sum would then require the three additional terms PE_{14}, PE_{24}, and PE_{34}, for a total of 1 + 2 + 3 = 6 PE_{ij} terms.

Proceeding in this fashion, you can see that calculating the total potential energy of a configuration of 10 charges would require 1 + 2 + 3 + … + 9 PE_{ij} terms in all, one for each pairing of charges in the configuration. There is a mathematical formula to determine this type of sum, which in this case has the value (9)(10)/2, or 45. Applying the formula for the PE of a pair of charges, 45 times over, and then adding up all the results, would be a lot of work! A computer might well prove handy to perform this tedious exercise.

25.6 - Interactive checkpoint: electric potential energy



The average distance between the proton and the electron in a hydrogen atom is 5.29×10^{-11} m. With the particles separated by this distance, the "Bohr radius," what is the electric potential energy of the configuration?

Answer:

$PE_e =$ ______ J

25.7 - Electric potential

Electric potential: The electric potential energy of a test charge at a given point in a field, divided by the test charge.

The electric potential of a location in an electric field is the measure of how much electric potential energy will be generated by placing a charge at that point. It is analogous to the concept of electric field. Just as the concept of electric field provides a way to determine how much electric force any charge will encounter at a given location, the concept of electric potential is used to determine how much potential energy placing the charge in a field will generate.

The term “electric potential” may seem confusing, since we are already using the related term “electric potential energy.” The formula in Equation 1 provides a direct way to calculate electric potential. The electric potential at a point in a field is calculated by placing a test charge in a field, determining the potential energy of the system created by introducing the charge at that location, and then dividing by the test charge.

The electric potential at any point is unique, which means it has only one value. This reflects the fact that the electrostatic force is conservative, and any configuration of charges has just one value for its potential energy. Electric potential has no direction, only a magnitude, making it a scalar. In Concept 1, you see a graph of the electric potential around a positive point charge, and in Concept 2, you see the graph of the electric potential around a negative point charge. Note how the electric potential increases near the positive point charge, and becomes increasingly negative near the negative point charge. This reflects how the PE_e of a positive test charge will increase near a positive charge, and take on larger negative values near a negative charge.

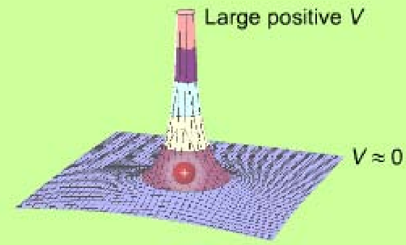
The graphs on the right provide metaphors for electric potential. The positive charge creates a potential peak, a mountain to be scaled, while the negative charge creates a potential well, a hole to be climbed out of.

You should remember that, just like electric fields, real potential peaks and wells exist around charges in three-dimensional space. They would require four dimensions to properly graph! The three-dimensional graphs in Concepts 1 and 2 only show the electric potential at locations lying in a plane around a point charge.

The equation for calculating the electric potential at locations around a point charge is shown in Equation 2. The equation can be derived from the equation for the PE_e of a pair of charges. The PE_e created by placing a test charge in the field is calculated, and then the PE_e is divided by the test charge, which means the test charge factor cancels out of the equation. Like the equation for potential energy from which it comes, this equation assumes there is zero potential energy when there is an infinite separation between the charges. This is the same as stating that the electric potential infinitely far away from a point charge is zero, as suggested by the graphs of the potential peak and the potential well.

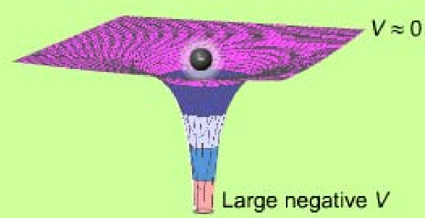
The unit of electric potential is the *volt*. A volt is defined as one joule per coulomb, or energy per unit charge. It is named after Alessandro Volta (1745-1827), the Italian inventor of the system that underlies the design of most batteries.

concept 1



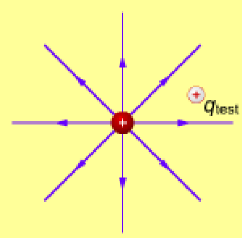
Electric potential
Reflects ability of fields to create *PE*
Calculated as electric *PE* / test charge

concept 2



Electric potential
Potential can be positive or negative

equation 1

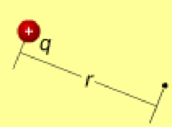


Electric potential of location in field

$$V = PE_e/q_{test}$$

V = electric potential
PE_e = electric potential energy
q_{test} = test charge
Units: volts (V)

equation 2

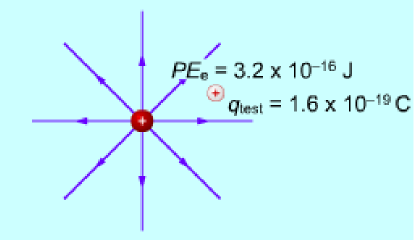


Electric potential around a point charge

$$V = \frac{kq}{r}$$

V = electric potential
k = Coulomb's constant
q = point charge
r = distance from charge

example 1



What is the electric potential at the location of the test charge?

$V = PE_e/q_{test}$

$V = (3.2\times10^{-16}\ \text{J})/(1.6\times10^{-19}\ \text{C})$

$V = 2.0\times10^{3}\ \text{V} = 2000\ \text{V}$

25.8 - Electric potential: multiple charges

The electric potentials from two or more charges can be added to determine their combined effect. We use the convention that the electric potential is zero at an infinite distance.

The two formulas in Equation 1 on the right describe the electric potential generated at a point in space by multiple point charges. The first is stated for three charges, and the second can be used with any number of charges. Both equations show that at any given point, the electric potential resulting from multiple charges is the sum of the electric potentials from each charge.

In Example 1 we show how to use the three-charge equation to determine the electric potential at a particular point. Three charges lie along the x axis. There is a charge of $1.00\ \mu\text{C}$ at position $x = -2.00\ \text{m}$, a charge of $2.00\ \mu\text{C}$ at $x = 1.00\ \text{m}$, and a charge of $-3.00\ \mu C$ at $x = 3.00\ \text{m}$. We sum the three individual electric potentials due to these charges at the origin to find the net electric potential there.

The solution to the example problem illustrates how to use a sum to determine an electric potential. It may also be helpful to see how the relationship between electric potential energy, work and force can be applied to calculate an electric potential. In order to do so, consider a point at an equal distance from two equal but opposite charges. The electric potential at this point is zero. This can be confirmed using the equation on the right, stated for two charges.

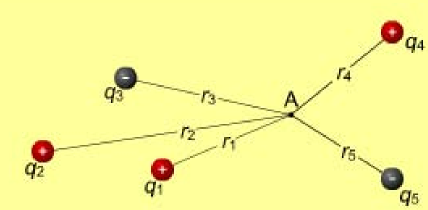
Now, we use the relationship between electric potential energy, work and force to arrive at the same conclusion. The electric potential energy of a third charge located at any equidistant point equals the negative of the work done by the field of the two stationary charges as the new charge moves in from infinity. Consider a charge moving in from infinity along the perpendicular bisector of the line segment between the two charges. Every point on such a line is equidistant from the two fixed charges. For example, in the illustration below, we show the third charge approaching along the y axis, with the fixed charges being at $x = -2.0$ and $x = 2.0$.

concept 1

Electric potential: multiple charges

Sum of potentials due to individual charges

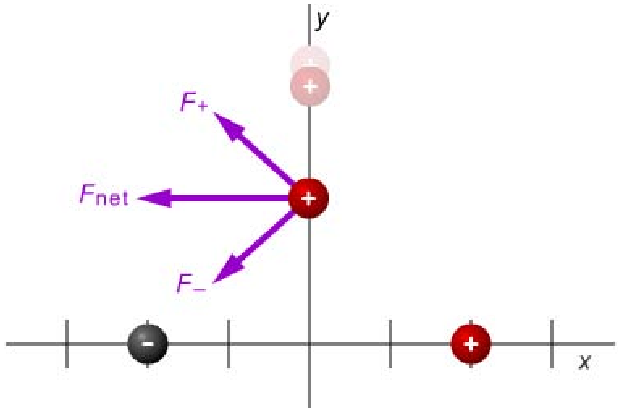
equation 1



Net potential is sum of potentials

For a system of 3 charges:

$$V = \frac{kq_1}{r_1} + \frac{kq_2}{r_2} + \frac{kq_3}{r_3}$$



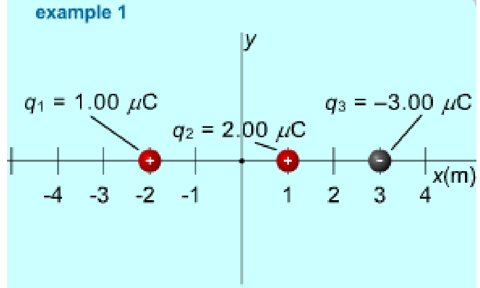
The net force exerted by the system on the test charge at any point on the y axis is **perpendicular** to its motion along the axis: The y components of the forces from the stationary charges cancel. Since there is no net force parallel to the path, only perpendicular to the path, this means there is no work, and no work means no change in potential energy away from its reference value of zero. Finally, zero potential energy means zero electric potential, which confirms the conclusion that the electric potential at any point equidistant from the charges equals zero.

V = potential at A
k = Coulomb's constant
q = charge
r = distance to charge

For a system of n charges:

$$V = \sum \frac{kq_i}{r_i} = k \sum \frac{q_i}{r_i}$$

example 1



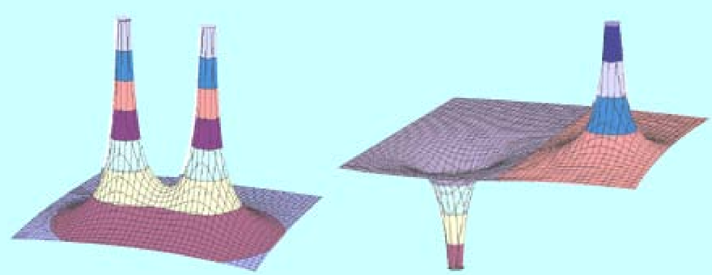
Find the electric potential at the position *x* = 0 m.

$$V = \frac{kq_1}{r_1} + \frac{kq_2}{r_2} + \frac{kq_3}{r_3}$$

$$V = k\left(\frac{1.00\mu\text{C}}{2.00\text{ m}} + \frac{2.00\mu\text{C}}{1.00\text{ m}} + \frac{-3.00\mu\text{C}}{3.00\text{ m}}\right)$$

$V = (8.99\times10^9\ \text{N·m}^2/\text{C}^2)(1.5\ \mu\text{C/m})$
$V = 1.35\times10^4$ V

25.9 - Sample problem: field vs. potential



Is it possible to have **E** = 0 and $V \neq 0$ at the same point?

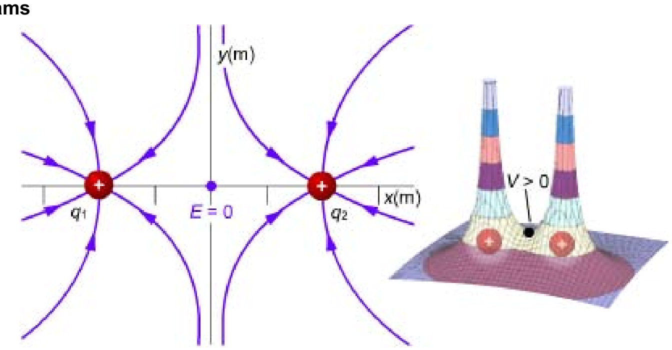
Is it possible to have **E** ≠ 0 and $V = 0$ at the same point?

Give examples to support your answers. Study the potential diagrams appearing to the left for some hints.

In fact, both situations are possible. We give two examples with diagrams and calculations.

Midway between equal charges. For this example we place two equal charges of +5.00 C at the points $x = \pm 2.00$ m on the x axis. We draw diagrams and perform calculations to find the values of $\mathbf{E}$ and V at the origin. We find that $\mathbf{E} = 0$ and $V \neq 0$.

Diagrams



Variables

	Left-hand	Right-hand
charge	$q_1 = +5.00$ C	$q_2 = +5.00$ C
distance from origin	$r_1 = 2.00$ m	$r_2 = 2.00$ m
field at origin	$\mathbf{E}_1$	$\mathbf{E}_2$
field strength at origin	E_1	E_2

Coulomb's constant	$k = 8.99\times10^9$ N·m²/C²
net field at origin	$\mathbf{E}_{\text{net}}$
potential at origin	V

What is the strategy?

1. Calculate the field vector at the origin due to each charge separately, and then add the vectors. The result will be the zero vector.
2. Calculate the total potential at the origin as the sum of the potentials due to each charge. The result will be greater than zero.

Physics principles and equations

The electric field magnitude at a distance r from a point charge q is given by

$$E = \frac{k|q|}{r^2}$$

The field is directed away from positive point charges, and toward negative point charges. By the principle of superposition, the net field due to two charges equals the vector sum of their individual fields.

The potential at a point in space due to two charges is given by

$$V = k\left(\frac{q_1}{r_1} + \frac{q_2}{r_2}\right)$$

Step-by-step solution

First we calculate the field $\mathbf{E}_{\text{net}}$ at the origin.

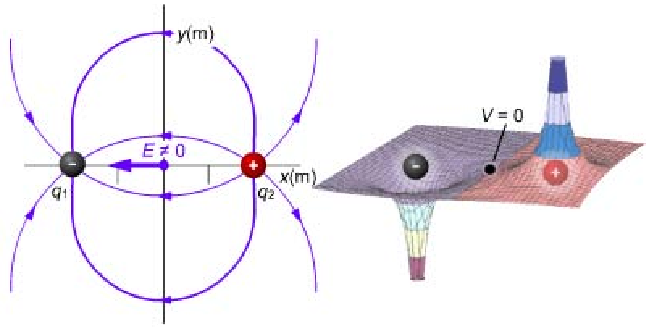
Step	Reason
1. $\mathbf{E}_{\text{net}} = \mathbf{E}_1 + \mathbf{E}_2$	superposition
2. $\mathbf{E}_1 = -\mathbf{E}_2$	inspection of field diagram
3. $\mathbf{E}_{\text{net}} = -\mathbf{E}_2 + \mathbf{E}_2 = 0$	substitute equation 2 into equation 1

Now we calculate the total potential V at the origin.

Step	Reason
4. $V = k\left(\frac{q_1}{r_1} + \frac{q_2}{r_2}\right)$	total potential
5. $V = (8.99\times10^9 \frac{\text{N}\cdot\text{m}^2}{\text{C}^2})\left(\frac{5.00\ \text{C}}{2.00\ \text{m}} + \frac{5.00\ \text{C}}{2.00\ \text{m}}\right)$	substitute values
6. $V = 4.50\times10^{10}$ J/C $V > 0$	evaluate

Midway between opposite charges. For this example we place two opposite charges of ±5.00 C at the points $x = \pm2.00$ m on the x axis. We draw diagrams and perform calculations to find the values of $\mathbf{E}$ and V at the origin. We find that $\mathbf{E} \neq 0$ and $V = 0$.

Diagrams



Variables

	Left-hand	Right-hand
charge	$q_1 = -5.00\ \text{C}$	$q_2 = +5.00\ \text{C}$
distance from origin	$r_1 = 2.00\ \text{m}$	$r_2 = 2.00\ \text{m}$

What is the strategy?

1. Calculate the field vector at the origin due to each charge separately, and then add the vectors. The result will be a nonzero vector.
2. Calculate the total potential at the origin as the sum of the potentials due to each charge. The result will be zero.

Step-by-step solution

First we calculate the field $\mathbf{E}_{\text{net}}$ at the origin.

Step	Reason
1. $\mathbf{E}_{\text{net}} = \mathbf{E}_1 + \mathbf{E}_2$	superposition
2. $\mathbf{E}_1 = \mathbf{E}_2$	inspection of field diagram
3. $\mathbf{E}_{\text{net}} = 2\mathbf{E}_2 \neq 0$	substitute equation 2 into equation 1

Now we calculate the total potential V at the origin.

Step	Reason
4. $V = k\left(\dfrac{q_1}{r_1} + \dfrac{q_2}{r_2}\right)$	total potential
5. $V = (8.99 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)\left(\dfrac{-5.00\ \text{C}}{2.00\ \text{m}} + \dfrac{5.00\ \text{C}}{2.00\ \text{m}}\right)$	substitute values
6. $V = 0$	evaluate

This last set of steps confirms a point made earlier: The electric potential midway between two equal but opposite charges − an *electric dipole* − equals zero.

25.10 - Spreadsheet: electric potential graphs

At the right, you see a three-dimensional graph of the electric potential for the points in a plane surrounding three charges. The graph was generated by a spreadsheet that you can access from this textbook if you are using a computer that can open Microsoft® Excel files. The electric potential at any point in the horizontal plane is indicated by the graph's distance above or below zero. Two positive charges of different magnitudes cause the "peaks" you see at the right, while the "well" is caused by a negative charge. The electric potential takes on extremely large positive or negative values at locations very near the positive or negative charges. We do not show the most extreme values because it would require a drastic change to the vertical scale of the graph.

The graph may remind you of very mountainous terrain. It serves as a map of electric potential, showing the locations of peaks, plains and wells. The electric PE of a test charge will increase as it "climbs up" the side of any peak. It requires positive work on the test charge to cause it to approach the underlying charge generating the peak. On the other hand, the PE of a test charge that is free to move will naturally "fall into" a well, as the test charge is attracted by the underlying negative charge.

Click here to launch the spreadsheet. You can change the amounts and locations of the charges, or add additional charges, and see the results. If the file does not open, on Windows click with the right mouse button and choose the save option. On the Macintosh, hold down the “control” key and click on the link, then choose the option to download the file.

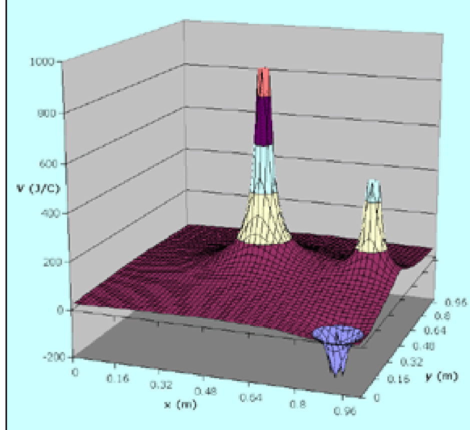
Click here to see a separate document that explains the programming of the spreadsheet in case you want to modify it.

Graphing Electric Potential

Charge Locations and Magnitudes

	x(m)	y(m)	Charge(μC)	Charge(C)
Charge 1	0.5	0.5	0.002	2.00E-09
Charge 2	0.9	0.1	-0.0005	-5.00E-10
Charge 3	0.9	0.5	0.001	1.00E-09

Electric Potential V vs. Location (x, y)



25.11 - Electric potential difference

Electric potential difference is the difference in electric potential between two points. For example, if point B has an electric potential of positive five volts, and point A has an electric potential of positive three volts, then the potential difference $V_B - V_A$ is two volts. You see this in Equation 1. Like electric potentials, electric potential differences are measured in volts.

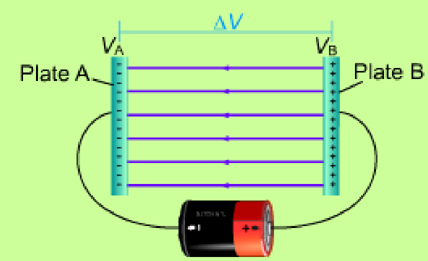
The potential difference between the terminals of the car battery is 12 V. Between two contacts of a household outlet, Δ*V* averages 120 V.

In the three illustrations to the right, you see two plates. They are both charged: The plate on the left (plate A) is negatively charged, and the plate on the right (plate B) is positively charged. In the example problem, we chose to make the electric potential of plate A negative 0.5 V and of plate B positive 1.0 V. The potential difference between the plates is 1.5 V.

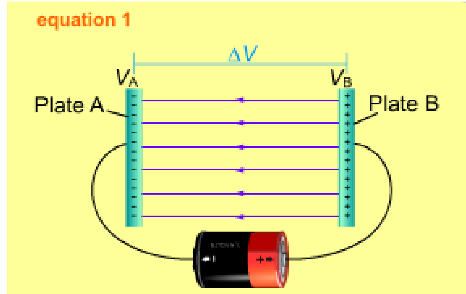
The electric potential values just cited for the plates depend on some choice of a reference point for zero potential. Often, in practical applications, the ground (the electrically neutral Earth) is defined as having zero electric potential, and all values are measured relative to this value. In an interactive problem in the introduction to this chapter, the midpoint between the plates was selected as the location of zero potential.

Also at the right, you see a common source of electric potential difference: a battery. Its two terminals have different electric potentials. Batteries are often described by the potential difference between their terminals. For instance, the potential difference between the terminals of a D cell is 1.5 V. This is often referred to as the battery’s voltage. In the configuration at the right, the potential difference between the battery’s terminals caused electrons to flow to plate A and away from plate B, thereby charging the plates.

concept 1



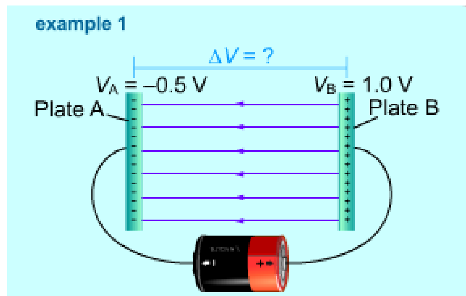
Electric potential difference
Difference in electric potential between two points



Potential difference

$\Delta V = V_B - V_A$

ΔV = electric potential difference
V_B = potential at point B
V_A = potential at point A
Units: volts (V)



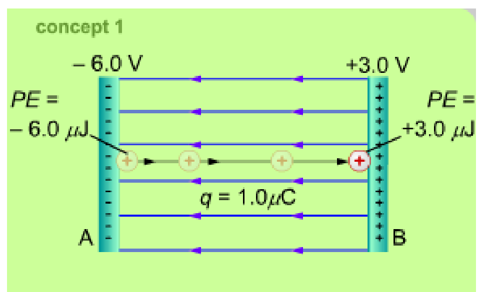
What is the potential difference between the plates?

$\Delta V = V_B - V_A$
$\Delta V = 1.0\text{ V} - (-0.5\text{ V})$
$\Delta V = 1.5\text{ V}$

25.12 - Potential difference, electric potential energy and work

At the right, you see two charged conducting plates. Plate A on the left is negatively charged, and plate B on the right is positively charged. The electric potential of plate A is −6.0 V, and that of plate B is +3.0 V. In this section, we will use this configuration to summarize the relationship of electric potential difference, electric potential energy and work.

The potential difference between the plates is +9.0 V (+3.0 V minus −6.0 V). How does this relate to potential energy and work? Consider a positive test charge with a charge of one microcoulomb that we cause to move between the plates. The charge is initially stationary at plate A. We move it to the positive plate B, and hold it still against this plate. The work we do changes only the charge's PE_e: It is stationary at its initial and final positions, so it has zero KE at both plates. The movement of the charge between the two plates is indicated in Concept 1.



Potential difference

Change in energy per unit charge

$V = PE_e/q$, and

$\Delta V = \Delta PE_e/q$

There is an electric field between the plates. It points from B to A, opposite to the direction we move the positive charge. Plate B repels the charge and plate A attracts it. This means we have to apply a force to move the charge from A to B. Since we are applying a force through a displacement, we are doing work on the system. The force we apply is in the same direction as the displacement of the test charge, so we are performing positive work. The force through a displacement is shown in Concept 2.

The (positive) work we do increases the electric potential energy of the positive test charge. It has more potential energy at B than at A.

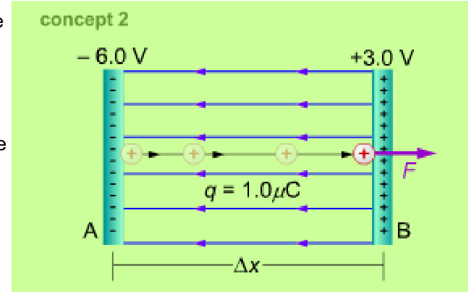
If you like, you can use a gravitational analogy. Consider the work required to roll a ball from an area of low gravitational potential, a valley, up an inclined ramp to an area of high gravitational potential, a mountaintop. Moving the ball uphill requires work, and the positive work done on the ball increases its potential energy. This resembles the positive test charge's trip from plate A to plate B. Conversely, if the ball rolls back down the hill (or the test charge is allowed to coast back to plate A), its potential energy decreases.

Now we discuss the same points, using equations. Electric potential can be calculated as the electric potential energy created by placing a test charge in a field, divided by the test charge. The electric potential **difference** is the **change** (difference) in electric potential energy divided by the test charge. These two equations are shown in Concept 1.

What causes this change in potential energy? The work done **on** the system. Taking the previous equation and replacing ΔPE with work W, we find that the electric potential difference is W/q, or work per unit charge. This is shown in Concept 2.

In all this discussion, a positive test charge has been used. This is the standard way to determine the electric potential. However, negative charges (like electrons) can move, too. When free to move, a negative charge will naturally move to an area of higher potential (the positive plate in this example). This will decrease the system's potential energy. With positive work, it can be forced to return to the negative plate, and this will increase the potential energy of the system. (As a form of shorthand, you may sometimes see this referred to as the PE of the charge, but it is important to remember that the charge only has PE because of the field caused by other charges.)

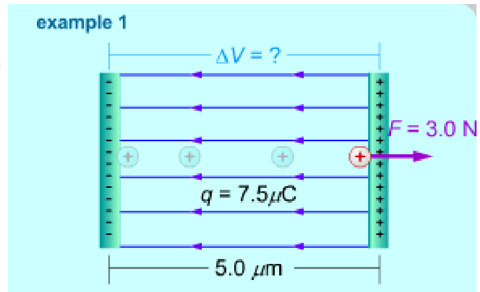
The *electron volt*, a common unit of energy in many areas of physics study, is based on the relationship of energy, charge and potential difference. One electron volt (eV) equals the change in the potential energy of an elementary charge (1.60×10^{-19} C) when it moves through a potential difference of one volt. We can use a rearrangement of the second equation in Concept 1, $\Delta PE_e = q\Delta V$, to calculate that an electron volt equals $1.602\ 18\times10^{-19}$ J. Note that the electron volt equals charge multiplied by electric potential, which equals electric potential **energy**. In spite of its name this unit refers to electric potential energy, not electric potential as volts do.



PE, work and potential difference

$W = \Delta PE_e$, so

$\Delta V = W/q$



What is the potential difference between the plates?

$W = F\Delta x$

$\Delta V = W/q$

$\Delta V = F\Delta x/q$

$\Delta V = (3.0\ \text{N})(5.0\ \mu\text{m})/(7.5\ \mu\text{C})$

$\Delta V = 2.0\ \text{V}$

25.13 - Interactive checkpoint: lead acid cell

A typical car battery consists of six lead-acid cells. Each cell has a potential difference of 2.05 V across it. The negative electrode is at a potential of −1.685 V, while the positive electrode is at a potential of +0.365 V. What is the change in electric potential energy for one electron that is moved from the positive to the negative electrode of one of the cells?

Answer:

$\Delta PE_e =$ ______ J

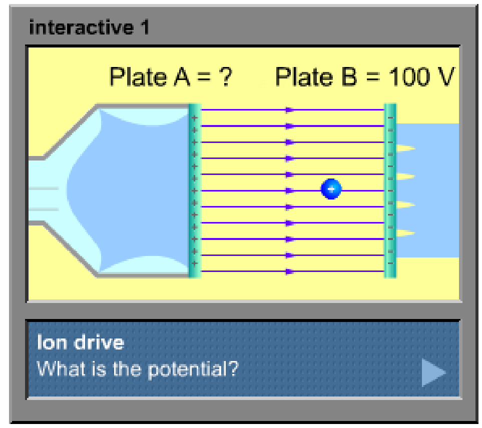
25.14 - Interactive problem: tune the rocket drive potential

Ion drives for spacecraft were introduced in another chapter. Two plates at different electric potentials are used to create an electric field that accelerates the particles in this ion drive. In the simulation, the electric potential of one plate is given; you are asked to determine what the potential of the other plate should be. You want an electric potential that will accelerate xenon ions of mass 2.20×10^{-25} kg and charge $+e$ from rest to a velocity of 30,000 m/s (3.00×10^{4} m/s) as they pass through the right-hand plate and leave the drive. The plates are separated by 1.00 m.

To solve this problem, you need to determine the potential difference between the plates. Use the relationship between electric potential, potential energy and charge to solve this problem. (Hint: The decrease in the ion's potential energy will equal its increase in kinetic energy.)

The plate on the right has an electric potential of 100 volts. Enter the electric potential of the other plate, to the nearest whole volt, in the space provided. Click on GO, and observe the exhaust velocity of the ions.

Interestingly, it turns out that the distance between the plates does not affect the answer to this problem. For an extra challenge, can you show why? If you have trouble answering these questions, refer to the definitions of electric potential and potential difference in terms of potential energy.



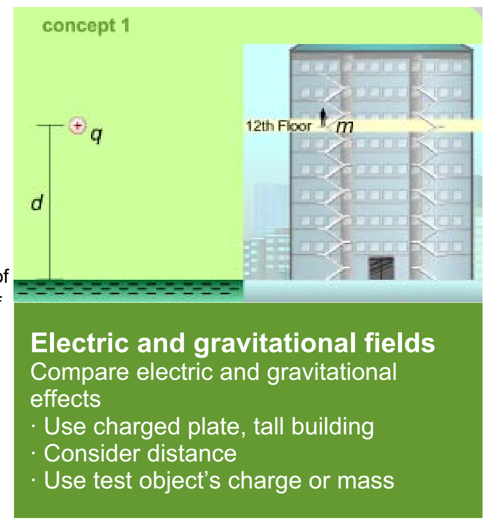
25.15 - A comparison of electric and gravitational fields

In this section, we compare the uniform electric field produced by a large negatively charged plate with the gravitational field created by the Earth. The goal is to use the gravitational field − one you experience all the time − to help you better understand the electric field.

In the gravitational field, the test objects in the illustrations to the right are people: Individuals of varying masses can walk up or down the stairs of a building to change their height above the Earth's surface. In the electric field, the test objects are positive test charges that can be farther from or closer to the charged plate. Their distance from the plate is indicated by d. In this discussion, the surface of the plate, and the Earth, are assumed to be points of zero potential energy.

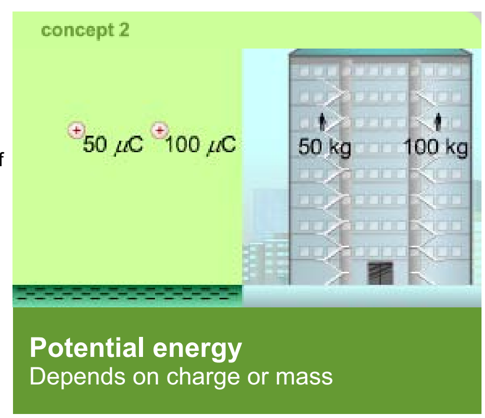
Points in space around Jupiter have more potential in its gravitational field than would points the same distance from the Earth in its gravitational field.

Let's now use the analogy. First, consider two people on the stairs. One person has a mass of 50 kg, the other 100 kg. Say they are both on the 12th floor of a building. Each person's potential energy is a function of his mass and height above the ground, and the acceleration due to Earth's gravitational field. Specific examples of this are illustrated in the Concept 2 diagram to the right: A 100 kg person on the 12th floor has twice as much gravitational potential energy as a 50 kg person on the same floor.



An analogous situation exists for a positive test charge above the center of the large negatively charged plate. The plate attracts the charge because it creates a field with the same downward orientation as the gravitational field. The electric potential energy of the charge is a function of its magnitude, its distance from the plate, and the strength of the field. In the Concept 2 diagram, two charges of different magnitudes are the same distance from the plate. The stronger charge will have more electric potential energy, just as the more massive person at a given height has more gravitational potential energy.

Now we will observe how position relates to potential energy. This is shown in the Concept 3 diagram. The same person is shown on the 12th and 6th floors. (To simplify the discussion, we use the European convention that the ground level is the "zeroth" floor, not the first, so that the 12th floor is twice as far above the ground as the 6th.) The person's gravitational potential energy is twice as much on the 12th floor as on the 6th floor. We also show two test charges of the same magnitude at different distances from the plate. The test charge at twice the distance from the plate has twice the potential energy.



Finally let's consider the role of the field. Potential energy is a function of the strength of the field-producing object. For example, if we moved the building to Jupiter, the same people on the same floors would have greater gravitational potential energy than they do on Earth. Jupiter's stronger gravitational field would account for the greater gravitational PE. If the plate had more charge and therefore a stronger field, test charges at the same positions would have greater electric potential energy. With a mass like Earth or with an electrically charged plate, you can think of fields as surrounding these entities, waiting to exert forces on nearby objects.

In both fields, the **work** done on an object equals its change in potential energy. For instance, if an elevator picks up a person at the 6th and drops him off at the 12th floor, the work done by the elevator equals the change in his potential energy. Similarly, if you exert a force to move a test charge farther from the plate, the work you do equals the change in the system's potential energy. (The person and charge are both stationary at the beginning and end of their journeys, so the work only changes their PE_e. Their KE at the beginning and end is zero.)

Now consider the concept of gravitational potential. As illustrated in the Concept 4 diagram, gravitational potential depends only on height above the Earth's surface. The gravitational potential is twice as great on the 12th floor as it is on the 6th floor.

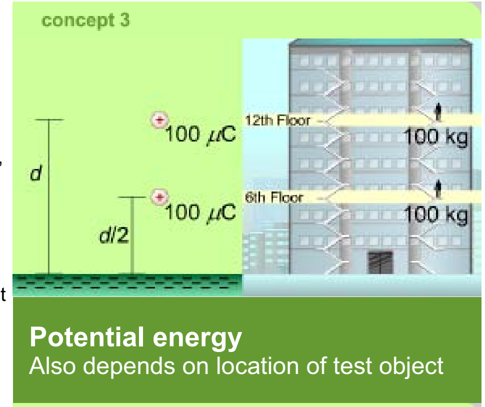
We can also compare positions 12 floors above the ground on Jupiter and on the Earth, two planets of different masses. The gravitational potential 12 floors above the ground is greater on Jupiter than it is on the Earth. To put it another way, any object you place on the 12th floor on Jupiter will have more potential energy than the same object would on the 12th floor on the Earth. This is because Jupiter's gravitational field is stronger than the Earth's. The important distinction between potential energy and potential is that the potential is independent of the test object. Jupiter's mass generates more gravitational potential at a given height above its surface than the Earth's mass does.

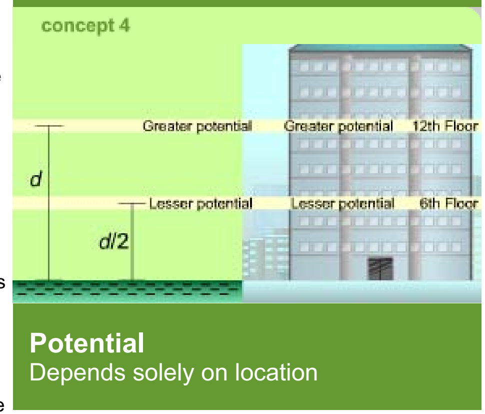
The concept of gravitational potential is analogous to electric potential. You see this also in the Concept 4 diagram. A location one centimeter above the plate has twice the electric potential of a location only half a centimeter above the plate. The distance above the plate determines the electric potential. It is independent of the strength of the test charge. Any charge, regardless of its magnitude, placed in an area of greater electric potential will create a system with more electric potential energy than when it is placed in an area of lower electric potential.

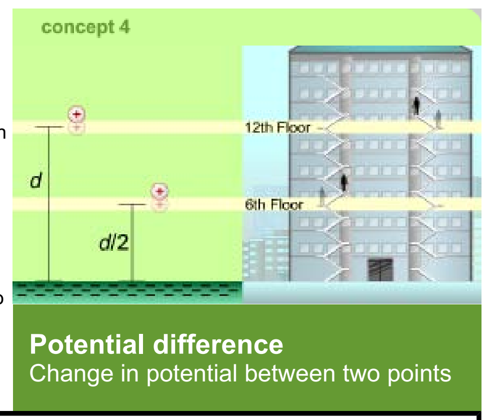
On the other hand, a location one centimeter from a highly charged plate with a stronger electric field has more potential than a location the same distance from a less charged plate. Again, this idea directly parallels the idea of gravitational potential. The statement for charged plates is true no matter what test charge you might use to assess the electric potential.

For the last idea, turn your attention to the illustration for the Concept 5 diagram. If you want to calculate the potential **difference** between the 6th and 7th floors, subtract the value of the potential at the 6th floor from its value at the 7th floor. Since we can treat the Earth's gravitational field as essentially uniform near its surface, the gravitational potential difference between the 6th and 7th floors is the same as it is between the 12th and 13th floors.

You should think of electric potential difference in a similar way. If you know the electric potentials at two points, you can subtract them to calculate the potential difference. With electric potential difference, the actual potentials at the points do not matter, only the change in potential between two points. The concept of electric potential difference is used frequently in everyday devices. For instance, if you are told you have a nine-volt battery, you are being told the difference between the terminals, **not** the electric potential of either terminal. This is analogous to how you might think if you were told to run up the stairs of a skyscraper. It does not matter whether you run from the 5th floor to the 15th, or from the 18th to the 28th. In either case, you have to run up 10 flights of stairs.

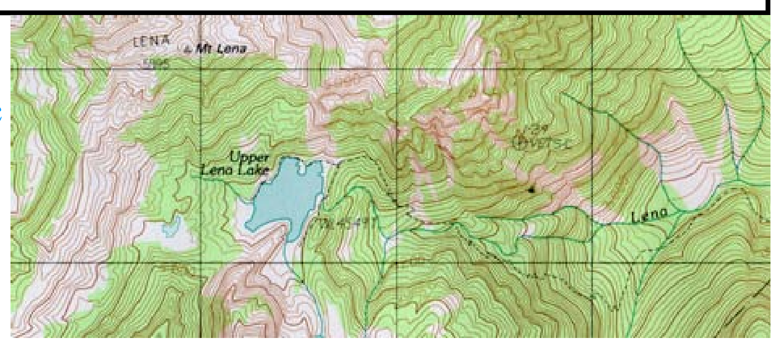






25.16 - Equipotential surfaces

Equipotential surface: A surface with the same electric potential everywhere.



The mountainous contour map above displays curves of constant altitude. Its "equi-altitude" curves are like equipotential surfaces in an electric field.

As its name indicates, an equipotential surface is one along which the electric potential of a field is everywhere the same. The boundary of a sphere centered about a point charge q is an equipotential surface. All the points on this surface are the same distance r from q. This means the electric potential, which in this case can be calculated with the formula kq/r, is the same at all points on the surface.

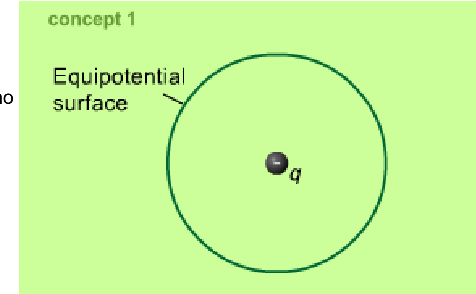
Such an equipotential surface is shown in the first diagram to the right. In the illustration, we represent the equipotential surface around q with a circle instead of a sphere for the sake of visual clarity.

No work is needed to move a charge from one resting place to another along an equipotential surface, because the potential energy neither increases nor decreases as the charge moves. If work were needed to change the charge's position, its electric potential energy would change, which means it would be in a location with a different electric potential.

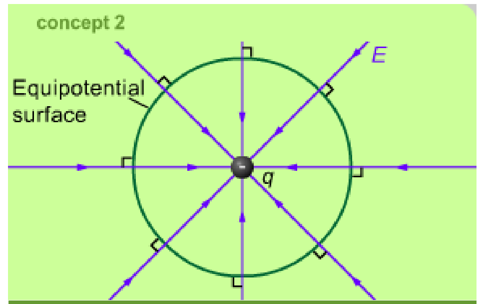
The second diagram shows a particular example of a relationship that holds true in general: Electric field lines that intersect an equipotential surface are always perpendicular to the surface where they intersect it.

This is true because any motion of a charged particle from one place to another along such a surface must be in a direction perpendicular to the force exerted by the field. Work equals the component of the force parallel to the motion, multiplied by the displacement. No parallel component means no work occurs. In turn, no work means no change in potential energy, and no change in potential energy means no change in potential.

You can explore equipotential surfaces in the simulations in the introduction to this chapter. If you move a test charge along an equipotential surface, you will see that its electric potential energy stays the same.



Equipotential surface
Surface with constant electrical potential
No work to move charge along surface



Equipotential surface
Electric field perpendicular to surface

25.17 - Electric potential and a uniform electric field

The concepts of electric potential and electric field are linked. In Concept 1, you see their relationship illustrated in the context of a uniform electric field. Plate A is negatively charged and plate B is positively charged. Assume that the plates are large enough compared to their spacing that the **electric field** between them is of uniform strength and direction, pointing from plate B toward plate A.

The **electric potential** between the plates increases in the opposite direction, from plate A to plate B. Since it takes positive work to push a test charge against the field from left to right, the potential energy of the charge increases as it moves from plate A to plate B. Since the strength of the field is everywhere the same, the electric potential increases at a constant rate as the charge moves from left to right. The potential's change per unit displacement parallel to the field is constant.

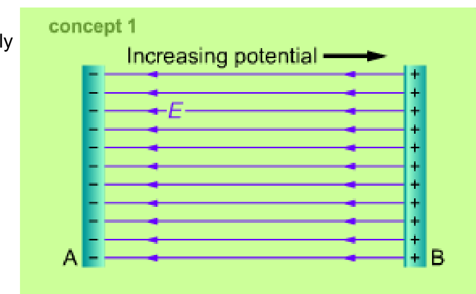
Another way of expressing the opposite orientations of the field and the direction of increasing potential is to say that the field is always directed from regions of higher potential to regions of lower potential.

We state the proportionality between the change of potential and the displacement as the first equation in Equation 1, and derive it below. The variable Δs measures a displacement parallel to the field lines. The negative sign in the equation reflects the fact that the potential difference is negative for a displacement $\Delta\mathbf{s}$ in the same direction as the field. Movement perpendicular to the field lines results in no change in electric potential. Again, we stress that the field must be uniform.

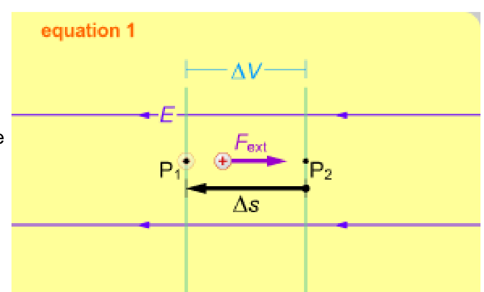
In many applications later in this book, we will only be interested in the magnitude of the potential difference between two points in a uniform field that are separated by a distance d parallel to the field, and we will write $\Delta V = Ed$.

The second equation in Equation 1 is just the first equation, solved for the electric field. It proves to be a very useful formulation. It shows how the electric field strength can be determined when the potential difference between two points in the field is known. This form of the equation is applied in the example problem below.

Derivation. We will show that for a uniform electric field $\mathbf{E}$, and a displacement $\Delta\mathbf{s}$ measured in the direction of the field, the potential difference between two points separated by Δs is given by the equation $\Delta V = -E\Delta s$. This is shown in the illustration of Equation 1.



Displacement parallel to uniform field
Potential V changes at constant rate
Field directed from higher to lower V



Potential difference in uniform field

$$\Delta V = -E\Delta s$$

Variables

uniform electric field	$\mathbf{E}$
two points in field	P_1, P_2
test charge	q_{test}
force	$\mathbf{F}$
work	W
distance parallel to field	Δs
potential difference between points	ΔV

We use the subscripts $_{field}$ and $_{ext}$ in the steps below to distinguish the forces exerted and the work done by the field and by an external force that moves a test charge against the field.

Strategy

1. Calculate the work done by an external force to move a test charge against the constant force of the field.
2. State the work done on the field in a different way, using the relationship of potential difference, work done on the field, and charge.
3. Set the two expressions for work equal to each other and simplify to get the desired equation.

Physics principles and equations

The work done to move a test charge from point P_1 to point P_2, opposite to the direction of $\Delta\mathbf{s}$, is $\mathbf{F}\cdot(-\Delta\mathbf{s})$. Since $\mathbf{F}$ is constant and parallel to the displacement between the points, this dot product can be written as the scalar product $F\Delta s \cos 0°$, or just

$$W = F\Delta s$$

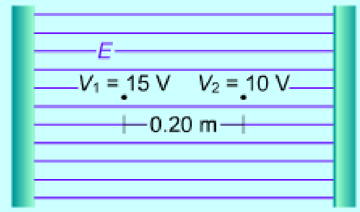
The force exerted on the test charge by the electric field is

$$F = qE$$

$$E = -\frac{\Delta V}{\Delta s}$$

ΔV = potential difference
E = electric field strength
Δs = displacement parallel to field

example 1



What are the magnitude and direction of this uniform electric field?

Field is directed to the right
$\Delta\mathbf{s}$ directed from left to right
ΔV = 10 V – 15 V = – 5.0 V

$$E = -\frac{\Delta V}{\Delta s}$$

E = – (– 5.0 V)/(0.20 m) = 25 N/C
$\mathbf{E}$ = 25 N/C to the right

A relationship stated and derived earlier for the relationship between potential difference and work done **on** a system is

$$\Delta V = W/q$$

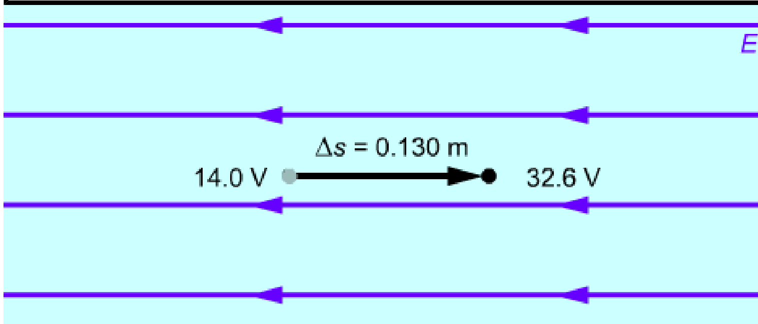
Step-by-step derivation

We find two expressions for the work W and set them equal to each other. We derive this using a positive charge being pushed against the field. The work done by the external force is positive in this case; work done by the field is negative.

Step	Reason
1. $W = F\Delta s$	definition of work
2. $F_{field} = q_{test}E$	electric field force on test charge
3. $F_{ext} = -q_{test}E$	external force on test charge
4. $W_{ext} = -q_{test}E\Delta s$	substitute equation 3 into equation 1
5. $W_{ext} = q_{test}\Delta V$	potential difference and work
6. $\Delta V = -E\Delta s$	equate steps 4 and 5 and simplify

Although we derived Equation 1 for the case of a positive charge being moved against a uniform field, the equation is equally true for negative charges, and for charges moving in either direction.

25.18 - Interactive checkpoint: an accelerated proton



In a uniform electric field, the potential changes from 14.0 V to 32.6 V in a displacement of 0.130 m (the direction is opposite to the field, as shown). What is the magnitude of the acceleration of a proton (charge 1.60×10^{-19} C and mass 1.67×10^{-27} kg) that is placed in this field?

Answer:

a = ______ m/s^2

25.19 - Gotchas

Electric potential energy and electric potential are the same. No, they are not even measured in the same units. Potential energy is a property of a specific configuration of charges, while electric potential is a property of a point in space. Electric potential is measured by assessing the potential energy with a test charge, and then dividing by the magnitude of that charge.

I quadrupled the strength of a ***test*** *charge, but I think the electric potential at the location is unchanged.* You are correct. Like an electric field, electric potential is independent of the test charge.

There is electric potential surrounding a single charge. Yes, there is. There is a field. Add a second charge, and you will have created a system with electric potential energy.

Where does an electron go when it is free to move: to a location of higher or lower potential? Toward a location of higher potential. Electrons tend to move toward positive charges and away from negative charges, which means they move toward regions of higher electric potential, minimizing the electric potential energy. A positive test charge (or proton) would do the opposite, moving toward a region of lower potential, which also minimizes the electric potential energy.

If I put a positive charge or a negative charge of the ***same magnitude*** *at a given point near a positive charge, in both cases I create a system with identical electric potential energy.* No. The magnitude of the potential energy will be the same, but the sign will be the opposite. Consider a positive charge. Placing another positive charge near it will require “pushing” the charge near it, increasing the potential energy of the system. Conversely, a negative charge will be attracted to the positive charge, and the closer it gets to the positive charge, the lower the potential energy of the system.

I calculated a negative potential difference between two points. Is that possible? Quite possible. If point A has an electric potential of 9.0 volts, and point B an electric potential of 6.0 volts, the difference $V_B - V_A$ is −3.0 volts. It is perhaps an unstated convention that when calculating a potential difference, V_A is subtracted from V_B.

Work, signs, potential energy: ***work done by field***. When the field generated by a system of particles does positive work, it reduces the **potential** energy of the system. For instance, if a charge attracts another charge, causing it to move closer, this is **positive** work done by the field, and it **reduces** the potential energy of the system. Think of gravity causing something to fall, and in doing so reducing the potential energy of the system. The total energy of the system stays the same: The field accelerates the object, increasing its KE.

Work, signs, potential energy: ***work done on the system by an external force***. Here, we are specifically discussing a charge that starts and stops at rest, so its initial and final KE are the same. An external force (like you) supplies the force. Positive work you do − work where the force lies along the path of the motion − increases the potential energy of the system. For instance, if you pull apart two charges with opposite signs, you increase their potential energy, just as when you lift a weight above the Earth’s surface, you increase the potential energy of the Earth-object system. Again, be careful with signs. You would have to force together two charges with like signs in order to do positive work and increase the energy of that configuration.

25.20 - Summary

Electric potential energy is conceptually similar to gravitational potential energy. It is determined by the configuration of a system of charges. As with gravitational potential energy, external positive work that is done **on** a system results in a positive change in the sum of the system's PE and KE. On the other hand, positive work done **by** the system on something outside the system results in a negative change in the sum of the system's PE and KE.

Since in this chapter we often assume that charges move from one stationary configuration to another, there is no change in KE due to work, and positive work done **on** a system increases its PE. Negative work done on a system, as when two opposite charges are moved closer together, decreases its PE.

The potential energy of a system of two charges is ordinarily considered to be zero when an infinite distance separates them.

Electric potential is a scalar quantity defined throughout an electric field in a region of three-dimensional space. It describes the electric potential energy that will be possessed by an electric test charge − or more properly by the system of charges − when the test charge is placed at various locations in the field. When a test charge is introduced, its electric potential energy equals the electric potential at its location, times its own charge value. The potential is measured in joules per coulomb, or volts (V), where 1 V = 1 J/C.

Equations

Potential energy of 2-charge system

$$PE_e = \frac{kq_1q_2}{r}$$

Electric potential

$$V = PE_e / q_{\text{test}}$$

$$V = \frac{kq}{r}$$

Potential difference in a uniform field

$$\Delta V = -E\Delta s$$

$$E = -\frac{\Delta V}{\Delta s}$$

The electric potential at a particular point is the sum of the potentials due to all the source charges present.

A dipole generates potential at the points in space around it. At large distances from the dipole, the potential is proportional to the magnitude of the dipole moment and inversely proportional to the square of the distance from the dipole.

Electric potential difference is more commonly measured and used than the electric potential itself. It is the difference in electric potential between two points. Batteries are categorized by the potential difference between their terminals, which is 1.5 V for commonly sold small batteries.

The electric potential difference between two points also equals the work required to move a test charge from one point to the other, divided by the value of the charge.

An equipotential surface in a field is a surface that has the same electric potential at all points. Because there is no potential difference between any two points in the surface, no work is required to move a charge along the surface. The electric field is perpendicular to such a surface at every point on the surface.

A vector called the potential gradient relates variations in potential to the local electric field. The gradient points in the direction of the fastest increase of the potential, and its magnitude equals the change in potential per unit distance in that direction.

Chapter 25 Problems

Chapter Assumptions

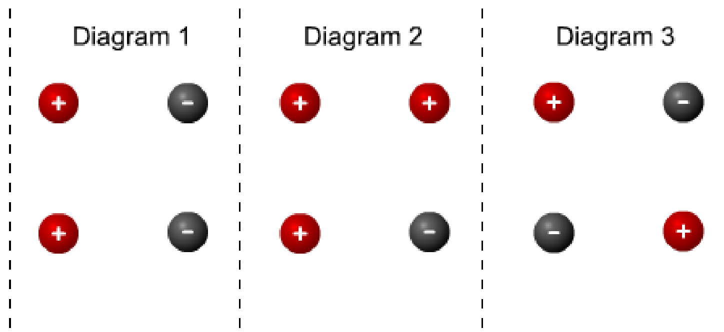
Unless stated otherwise, the reference configuration for zero electric potential and zero electric potential energy is one in which there is infinite separation between charges.

Conceptual Problems

C.1 A positively charged particle and a negatively charged particle are located near one another in space. Their masses, the magnitudes of their charges, and the distance between them are such that their gravitational potential energy equals their electric potential energy. (The reference configuration for zero gravitational *PE*, like that for zero electric *PE*, is an infinite separation of the particles.) (a) If the distance between the particles is halved, what happens to the electric potential energy of the configuration? (b) Is the gravitational potential energy of the new configuration still equal to the electric potential energy? (c) Are they still equal if one of the charges reverses in sign but keeps the same magnitude?

(a) i. It halves
ii. It doubles
iii. It stays the same
(b) ○ Yes ○ No
(c) ○ Yes ○ No

C.2 The following three configurations are composed of charges of the same magnitude, but differing signs, arranged at the corners of a square. (a) Which configuration has the greatest electric potential energy? (b) Which configuration has the most negative electric potential energy?



(a) i. 1
ii. 2
iii. 3
(b) i. 1
ii. 2
iii. 3

C.3 Explain the difference between electric potential energy and electric potential. How are the two related?

C.4 Electric field strength can be measured in N/C, but in view of the equation $E = |\Delta V/\Delta s|$, it is often written with the units V/m. Using only metric units and their definitions, show that 1 N/C = 1 V/m.

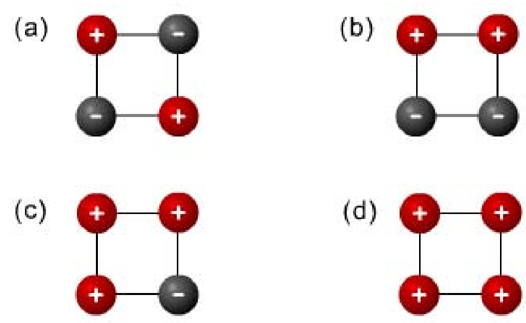
C.5 There is a charged ring of constant radius lying in the *yz* plane, centered at the origin. Suppose the charge can move, so the charge density is not necessarily uniform around the ring. Consider a point *P* on the *x* axis, the axis of the ring. (a) Does the potential at *P* vary as the distribution of charge on the ring varies? Why or why not? (b) Does the electric field at *P* vary as the distribution of charge on the ring varies? Why or why not?

(a) ○ Yes ○ No
(b) ○ Yes ○ No

C.6 (a) Two equal positive charges are to be placed on two corners of a square so that the electric potential at the other two corners is equal. Check all the configurations of the charges that will have this result. (b) A positive charge and a negative charge of equal magnitude are to be placed on two corners of a square so that the electric potential at the other two corners is equal. Check all the configurations of the charges that will have this result.

(a) ☐ Opposite corners
☐ Adjacent corners
(b) ☐ Opposite corners
☐ Adjacent corners

C.7 You have a selection of protons and electrons. You want to place them at the corners of a square so that the electric field and the electric potential at the center of the square are both zero. Check off all of the shown configurations that will produce this result.



- ☐ a
- ☐ b
- ☐ c
- ☐ d

C.8 Points A and B are arbitrarily located in a uniform electric field that extends throughout all of space. Your friend claims, "Whether you stand at A or at B, your surroundings are identical. In other words, the field looks the same from every point in space, so no point is different from any other. Therefore, the electric potential should be the same everywhere in space, no matter where you are. There can be no potential difference between any two points A and B." The conclusion is wrong, so what is the flaw in your friend's argument?

C.9 (a) Can equipotential surfaces with different electric potentials ever intersect each other? (b) Why or why not? (c) Do any two equipotential surfaces have to be a constant distance apart? Explain.

(a) ○ Yes ○ No
(b) Submit your answer on paper
(c) ○ Yes ○ No

Section Problems

Section 0 - Introduction

0.1 Using the simulation in the first interactive problem in this section, answer the following questions. (a) What is the sign of the *PE*? (b) Is the magnitude of the potential energy greater when the test charge is closer to, or farther away from the stationary charge?

(a) i. Always positive
ii. Always negative
iii. Both positive and negative

(b) ○ Closer to ○ Farther away

0.2 Using the simulation in the second interactive problem in this section, answer the following questions. (a) What is the sign of the *PE*? (b) As the distance between the charges increases, does the *PE* increase or decrease? (c) In what type of path should you move the charge in order to keep its *PE* constant?

(a) i. Always positive
ii. Always negative
iii. Both positive and negative

(b) i. Increases
ii. Decreases
iii. Stays the same

(c) i. A straight line
ii. A circle
iii. A parabola
iv. A square

0.3 Using the last interactive in this section answer the following questions. (a) How does the potential energy change as you move the test charge away from the negative plate and toward the positive plate? (b) In what type of path should you move the test charge in order to keep its *PE* constant?

(a) i. It increases
ii. It decreases
iii. It stays the same

(b) i. A circle
ii. A horizontal line
iii. A vertical line
iv. A square

Section 1 - Electric potential energy

1.1 A system consisting of two positive particles is altered by an outside force so that its potential energy increases by 5.4 μJ. There is no change in kinetic energy between the initial and final states of the system. (a) What is the work done by the outside force? (b) What is the work done by the electric force? (c) Are the particles closer together or farther apart after the system's rearrangement?

(a) _________ μ J
(b) _________ μ J
(c) i. Closer together
ii. Farther apart
iii. No change

1.2 A system consisting of one positively charged particle and one negatively charged particle is altered by an outside force so that its potential energy increases by 6.25 mJ. There is no change in kinetic energy between the initial and final states of the system. (a) What is the work done by the outside force? (b) What is the work done by the electric force? (c) Are the particles closer together or farther apart after the system's rearrangement?

(a) _________ mJ
(b) _________ mJ
(c) i. Closer together
ii. Farther apart
iii. No change

1.3 A system consisting of two positive particles is altered by an outside force so that its potential energy decreases by 52.3 J. There is no change in kinetic energy between the initial and final states of the system. (a) What is the work done by the outside force? (b) What is the work done by the electric force? (c) Are the particles closer together or farther apart after the system's rearrangement?

(a) _________ J
(b) _________ J
(c) i. Closer together
ii. Farther apart
iii. No change

Section 2 - Sample problem: electric potential energy

2.1 A system consists of a uniform electric field of 73.5 N/C, directed in the positive *x* direction, together with a particle having a positive charge of 3.46×10^{-5} C. The particle is moved from the location x = 3.50 m to the location x = 1.50 m in the field, and it is stationary before and after it is moved. (a) What is the work done **on** the system by the external force? (b) What is the work done **by** the field on the particle? (c) What is the change in the potential energy of the system?

(a) _________ J
(b) _________ J
(c) _________ J

2.2 A system consists of a uniform electric field of 42.6 N/C, directed in the positive *x* direction, together with a particle having a positive charge of 7.87×10^{-6} C. The particle is moved from the location x = 1.50 m to the location x = 3.50 m in the field, and it is stationary before and after it is moved. (a) What is the work done **on** the system by the external force? (b) What is the work done **by** the field on the particle? (c) What is the change in the potential energy of the system?

(a) _________ J
(b) _________ J
(c) _________ J

2.3 A system consists of a uniform electric field of 2.50 N/C, directed in the positive *x* direction, together with a particle having a negative charge of 5.50 mC. The particle is moved from the location x = 1.50 m to the location x = 3.50 m in the field, and it is stationary before and after it is moved. (a) What is the work done **on** the system by the external force? (b) What is the work done **by** the field on the particle? (c) What is the change in the potential energy of the system?

(a) _________ J
(b) _________ J
(c) _________ J

2.4 This problem explores the relationship between electric potential energy and work done **on** a system (by a non-conservative force) when there is a change in the *KE* of the system after the work is done. A system consists of a uniform electric field of 73.5 N/C, directed in the positive *x* direction, together with a particle having a positive charge of 3.46×10^{-5} C and a mass of 2.42×10^{-9} kg. An external force moves the initially stationary particle from the location x = 3.50 m to the location x = 1.50 m in the field, and when it stops "pushing" the particle is moving at 346 m/s to the left. (a) What is the change in the kinetic energy of the system? (b) What is the change in the potential energy of the system? (c) What is the work done **on** the system by the external force? (d) Is the change in the potential energy of the system still equal to the work done on the system?

(a) __________ J
(b) __________ J
(c) __________ J
(d) ○ Yes ○ No

2.5 This problem explores the relationship between potential energy and work done **by** a system (by a conservative force) when there is a change in the *KE* of the particle after the work is done. The system consists of a uniform electric field of 42.6 N/C, directed in the positive *x* direction, together with a particle with a positive charge of 7.87×10^{-6} C. As part of an experiment, the field is allowed to accelerate the particle starting from rest at x = 1.50 m. The experiment ends when the particle passes x = 3.50 m, at which time it is moving to the right. (a) What is the change in the kinetic energy of the system? (b) What is the change in the potential energy of the system? (c) What is the work done **by** the field on the particle? (d) Is the change in the potential energy of the system still equal to the negative of the work done by the conservative force?

(a) __________ J
(b) __________ J
(c) __________ J
(d) ○ Yes ○ No

Section 3 - Electric potential energy and work

3.1 Two positive charges of 3.6 μC and 2.5 μC are separated by a distance of 0.031 m. What is the electric potential energy of this configuration?

__________ J

3.2 A 6.3 μC point charge and a −1.7 μC point charge are separated by a distance of 0.00032 m. What is the electric potential energy of this system?

__________ J

3.3 The nucleus of a helium atom consists of two protons, each with a charge of +*e*, and two electrically neutral neutrons. The distance between the repelling protons is, on the average, about 1×10^{-15} m. They are constrained there by a force called, appropriately enough, the strong force. For each of the following questions, it is sufficient to report your answer with just one significant digit, multiplied by a power of 10. (a) In a classical, non-quantum model of the helium nucleus, what is the average strength of the repulsive force between the two protons? (b) What is the average electric potential energy of the nucleus?

(a) __________ N
(b) __________ J

3.4 The Bohr model of the hydrogen atom describes it as a proton (charge +*e*) orbited by a single electron (−*e*). When the atom is in its lowest energy state, the orbital radius of the electron is 5.29×10^{-11} m, a distance called the *Bohr radius*. What is the PE_e of a simple system consisting of a proton and an electron separated by that distance?

__________ J

3.5 The gravitational potential energy of all the water stored behind a very large dam is 1.2×10^{16} J, using the base of the dam as the reference point for zero *PE*. Imagine you could store the same amount of potential energy in a system consisting of two particles, each with a charge of 0.0020 C. How far apart should you place the charges?

__________ m

3.6 Two small iron spheres each have a charge of 0.250 C. (a) How much work does it take to move the spheres from a stationary configuration 2.00 m apart to another stationary configuration 1.00 m apart? (b) How much work does it take to move the spheres from a stationary configuration 12.0 m apart to another stationary configuration 11.0 m apart? (c) How much work does it take to move the spheres from a stationary configuration 22.0 m apart to another stationary configuration 21.0 m apart?

(a) __________ J
(b) __________ J
(c) __________ J

3.7 Two identical insulating spheres are separated by 1.25 m in a stationary configuration. The spheres carry opposite charges of magnitude 3.75 mC which are uniformly distributed in the spheres. Each ball has a radius 1.38 cm, and a mass of 2.07×10^{-2} kg. In a space-station experiment where gravity can be ignored, the spheres are allowed to drift together and "crash". What is their total kinetic energy just when they hit?

__________ J

3.8 Two infinitesimally small point charges of +2.0 μC each are initially stationary and 5.0 m apart. (a) How much work is required to bring them to a stationary position 2.0 m apart? (b) 2.0 mm apart? (c) With a finite amount of energy, is it possible to push them together?

(a) __________ J
(b) __________ J
(c) ○ Yes ○ No

Section 4 - Electric potential energy: multiple point charges

4.1 Choose the correct simplified formula for the potential energy of each of the following two systems.
(a) A system of three charges q_1, q_2, and q_3 that lie at the vertices of an equilateral triangle of edge r.

1) $\dfrac{3kq^2}{r}$

2) $\dfrac{k}{r}(q_1q_2+q_1q_3+q_2q_3)$

3) $\dfrac{k^3}{r^3}(q_1q_2+q_1q_3+q_2q_3)$

(b) A system of three equal charges q that lie at the vertices of an equilateral triangle of edge r.

1) $\dfrac{3kq^2}{r}$

2) $\dfrac{3kq}{r}$

3) $\dfrac{kq^2}{r^2}$

(a) i. 1
ii. 2
iii. 3
(b) i. 1
ii. 2
iii. 3

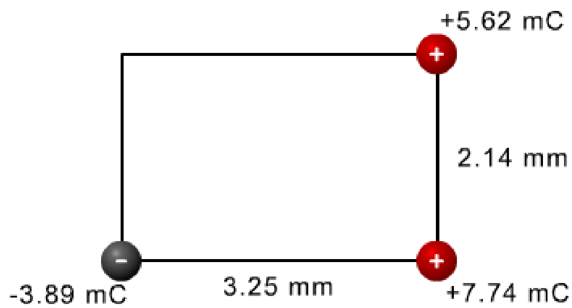
4.2 Four charges of magnitude +1.0 μC, +2.0 μC, +3.0 μC, and +4.0 μC lie at the vertices of a tetrahedron (three-sided pyramid). The distance of each charge from every other charge is exactly 1.0 cm. What is the potential energy of this configuration? State your answer to the nearest joule.

__________ J

Section 5 - Sample problem: multiple point charges

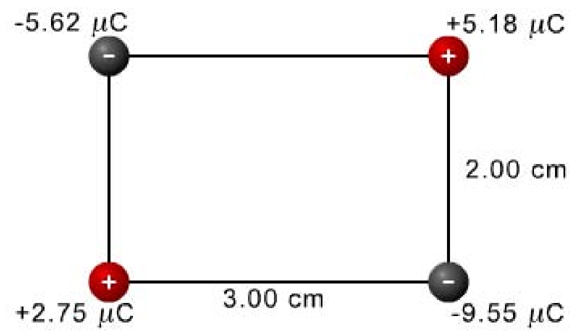
5.1 (a) Find the electric potential energy of the configuration of charges shown in the diagram. (b) Find the electric potential energy of the system if the charges of +7.74 mC and −3.89 mC exchange positions.

(a) __________ J
(b) __________ J



5.2 Find the electric potential energy of the configuration of charges shown in the diagram.

__________ J



5.3 A charge of +3.00 μC is placed at the point (0 m, 0 m, 0 m); a charge of −4.50 μC is placed at the point (2.00 m, −5.00 m, 7.00 m); finally, a charge of −1.00 μC is placed at the point (1.00 m, 1.00 m, 1.00 m). (a) What is the electric potential energy of this configuration? (b) If a neutron (a particle with the mass of the proton but no electric charge) is brought in from infinity and placed at the point (−3.00 m, 2.00 m, 5.50 m), how does the electric potential energy of the configuration change?

(a) __________ J
(b) i. It increases
ii. It decreases
iii. It stays the same

Section 7 - Electric potential

7.1 (a) Calculate the potential 1.00 m from an isolated proton. (b) Calculate the potential 1.00 m from an isolated electron. (c) Which potential is greater? (d) Which potential has a greater magnitude?

(a) __________ V
(b) __________ V
(c) i. Near the proton
ii. Near the electron
iii. Potentials are equal
(d) i. Near the proton
ii. Near the electron
iii. Magnitudes are equal

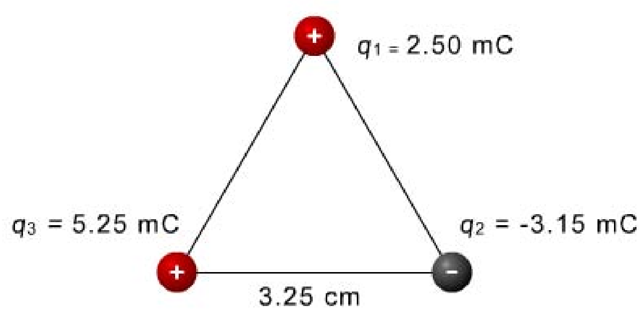
7.2 In a science fiction story, a microscopic black hole is given an enormous positive charge by firing an un-neutralized ion drive exhaust (consisting of positively charged xenon ions) into it for six months. The idea behind the charging process is to be able to confine and manipulate this dangerous object with powerful electric fields. Suppose the charge on the black hole is 5740 C. At what distance from it is the electric potential equal to 1.09e+3 V?

__________ m

7.3 A positively charged, small region in a thundercloud creates an electric potential of 67.6 mV at ground level, 321 m below. What is the charge of the region?

__________ C

7.4 The three charges in the diagram lie at the vertices of an equilateral triangle. (a) What is the electric **potential energy** of the configuration of three charges? (b) What is the electric **potential** due to q_2 and q_3 at the position of charge q_1? (c) What is the electric potential due to q_1 and q_3 at the position of charge q_2? (c) What is the electric potential due to q_1 and q_2 at the position of charge q_3?



(a) __________ J
(b) __________ V
(c) __________ V
(d) __________ V

7.5 Six charges, including both positive and negative ones, are placed in a configuration so that the electric potential at a certain location in space is 34.7 V. A seventh charge of −9.40e−6 C is brought in from infinity and deposited at this location. What is the change in electric potential energy of the resulting configuration?

__________ J

7.6 A nonconducting sphere of radius R = 5.00 cm contains a total charge of 4.25×10^{-8} C, uniformly distributed throughout its volume. What is the potential due to this charge distribution at a point at distance r (measured in meters) from the center of the sphere, where $r > R$? Express your answer in terms of r. (Hint: According to a version of the shell theorem, the force exerted by the sphere on a test charge at any external point is the same as if all the charge of the sphere were concentrated at its center.)

- ○ 382/(0.05+r)
- ○ 382/r
- ○ 382/(r-0.05)

7.7 An insulating sphere of radius R = 5.00 mm contains a total charge of 4.25×10^{-12} C, uniformly distributed over its surface (and fixed in place). An electron starts from rest at a distance of 9.00 mm from the surface of the sphere, and accelerates straight toward it. How fast is the electron moving when it crashes into the surface of the sphere? (Hint: According to a version of the shell theorem, the force exerted by the sphere on a test charge at any external point is the same as if all the charge of the sphere were concentrated at its center.)

__________ m/s

7.8 Beginning in 1906, Ernest Rutherford carried out a famous series of experiments at McGill University in which he bombarded thin gold foil with alpha particles having a mass of 6.68×10^{-27} kg and a charge of $+2e$. Most of the particles whizzed right through the foil, some were slightly deflected, and a very few were bounced (or "scattered") straight back. These results led him to speculate that every gold atom consisted of a tiny, compact, positively charged "nucleus" surrounded by a relatively huge volume of practically empty space, whose extent was defined by the orbits of its circling electrons. If an alpha particle is fired from a very distant source with an initial speed of 2.25×10^{7} m/s, straight at the nucleus of a gold atom, having a charge of $+79e$, how close does the particle get to the nucleus before it stops and gets scattered backwards?

__________ m

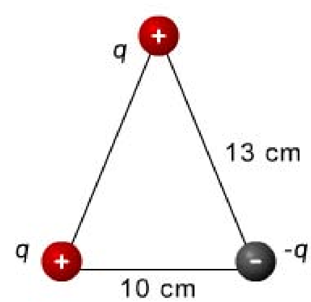
Section 8 - Electric potential: multiple charges

8.1 Two charged particles, $q_1 = -2.50\times10^{-3}$ C and q_2 = + 3.50e−3 C, are separated by a distance of 5.00 cm. (a) What is the electric potential energy of this configuration? (b) What is the electric potential at a point midway between the particles?

(a) __________ J
(b) __________ V

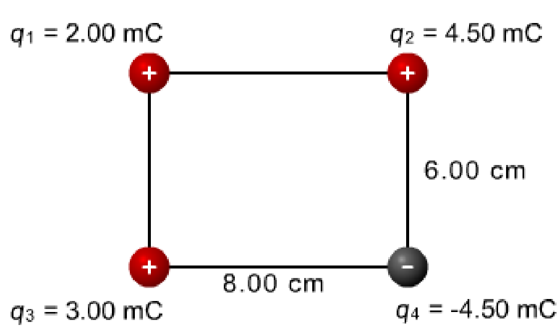
8.2 Three charges are arranged on the vertices of an isosceles triangle as shown in the diagram. What is the electric potential at the midpoint of the base of the triangle, in terms of q?

__________ q



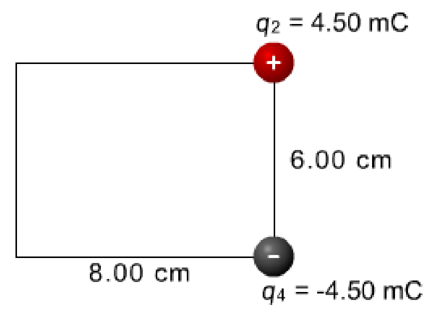
8.3 Four charges are arranged at the corners of a rectangle as shown in the diagram. (a) What is the electric potential at the center of the rectangle? (b) What is the electric potential there if the charges q_2 and q_4 are interchanged?

(a) __________ V
(b) __________ V



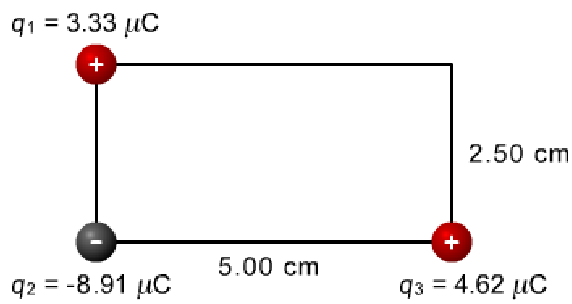
8.4 Two charges, q_2 and q_4, are arranged at the corners of a rectangle as shown in the diagram. (a) What is the electric potential at the center of the rectangle? (b) What is the electric potential there if the charges q_2 and q_4 are interchanged?

(a) ________ V
(b) ________ V



8.5 Three charges are arranged at the corners of a rectangle as shown in the diagram. What is the electric potential at the corner of the rectangle that does not have a charge?

________ V

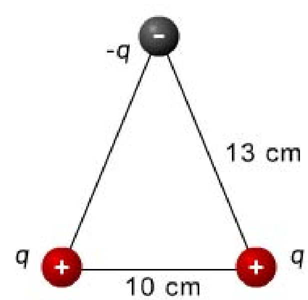


8.6 (a) Three charges q_1, q_2, and q_3 are located on the circumference of a circle of radius *r*. What is the potential due to these charges at the center of the circle? Express your answer as a formula involving the charges, the radius, and Coulomb's constant. (b) Write a similar formula for *n* charges, q_1, q_2, ... q_n, on the circumference of the circle. (c) A circular ring of wire, with radius 4.75 cm, carries a total charge of 23.5 μC. What is the potential due to this ring at its center?

(a) ○ $k(q_1+q_2+q_3)/r$
○ $k(q_1+q_2+q_3)/(3r)$
○ $3k(q_1+q_2+q_3)/r$
(b) Submit answer on paper.
(c) ________ V

8.7 Three charges are arranged on the vertices of an isosceles triangle as shown in the diagram. What is the electric potential at the midpoint of the base of the triangle, in terms of *q*?

________ q



Section 11 - Electric potential difference

11.1 The negative (black) terminal of a 12 V automobile battery is grounded to the frame of a car. If we arbitrarily take the automobile frame to be at an electric potential of 20 V, what is the electric potential of the positive (red) terminal of the battery?

________ V

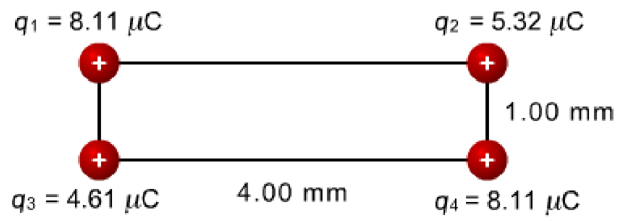
Section 12 - Potential difference, electric potential energy and work

12.1 An electron with charge -1.60×10^{-19} C moves a distance of 3.17 mm in the negative *x* direction in a uniform electric field of 125 N/C that points in the positive *x* direction. (a) How much work does the field do on the electron? (b) What is the change in the potential energy of the electron? (c) Through what potential difference does the electron move?

(a) ________ J
(b) ________ J
(c) ________ V

12.2 Four charges are arranged at the corners of a rectangle as shown in the diagram. How much work needs to be done to move the charge q_2 to infinity?

__________ J



12.3 The electric potential difference across the cell membrane of a human cell is 0.080 V, with the higher potential being outside the cell. In one cellular process a sodium ion Na^+, with charge $+e$, is "pumped" through a channel in the membrane to be ejected from the cell. (a) How much work must be done by the cell to eject the ion? (b) A candy bar contains 270 Calories, or 1.13×10^6 J of food energy. How many sodium ions could the cells of your body eject using the energy obtained in one candy bar?

(a) __________ J
(b) __________ ions

12.4 A planet far, far away was impacted by a one-kilogram ball of pure "protonium" containing 6.00×10^{26} protons, and nothing else. This excess charge eventually distributed itself evenly through the planet, which had previously been electrically neutral. (a) If the radius of the planet, like Earth's, is 6.37×10^6 m, what is the magnitude of the electric field at the planet's surface? (Hint: According to a version of the shell theorem, the force exerted by a spherically symmetric charge distribution on a test charge at any external point is the same as if all the charge were concentrated at its center.) (b) What is the potential differnce between a location 500 m above the surface of the planet and a location at ground level?

(a) __________ N/C
(b) __________ V

12.5 A certain 4.8-volt cellular telephone battery has a capacity of 600 mAh. (One mAh equals 3.6 coulombs.) When the battery's charge is too low to run the phone any longer (assume it is zero), the battery must be recharged. (a) How much work is done by the recharger? (b) If electricity costs you 9.0 cents per kWh, how much does one recharge cost you?

(a) __________ J
(b) $ __________

12.6 In Ernest Rutherford's scattering experiment, described in detail in an earlier homework problem in this chapter, alpha particles (with a mass of 6.68×10^{-27} kg and a charge of 3.20×10^{-19} C) come from a few milligrams of radium (to be very precise, its decay product radon). They leave with an initial speed of 2.25×10^{7} m/s before being directed to strike a thin sheet of gold foil. What is the change in electric potential (watch the sign of your answer) that would be needed to accelerate an alpha particle, starting at rest, up to this speed? Ignore relativistic effects.

__________ V

12.7 A charged particle of mass 8.39×10^{-4} kg starts from rest and accelerates through a potential difference of +25,000 V to reach a speed of 1.26e+3 m/s. What is the charge on this particle?

__________ C

12.8 Two parallel charged plates are held at a potential difference ΔV. Point A is on the plate with the lower potential, and point B is directly across from it on the other plate. An electron (of mass 9.11×10^{-31} kg) and a proton (1.67×10^{-27} kg) are released simultaneously from points A and B, respectively. The electron accelerates toward B and the proton accelerates toward A. Assume they do not collide or interact with each other. (a) Which particle reaches the opposite plate first? (b) How many times longer does it take the slower particle to complete this "drag race"?

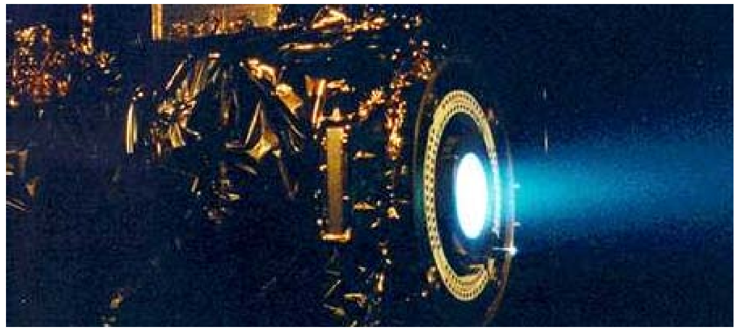
(a)
i. The electron
ii. The proton
iii. They take the same amount of time

(b) __________ times longer

Section 14 - Interactive problem: tune the rocket drive potential

14.1 Xenon ions with a mass of 2.20×10^{-25} kg and charge +*e* are accelerated between the plates of an ion drive rocket engine to an exhaust velocity of 43,200 m/s. (a) What is the potential difference between the plates? (b) If the potential difference is doubled, what is the velocity of the engine exhaust?

(a) __________ V
(b) __________ m/s



14.2 In a typical television or in an older computer monitor's cathode ray tube (CRT), electrons are accelerated from rest through a potential difference of 2.5×10^4 V, steered by magnetic fields, and finally strike particular spots on the screen at the front of the tube to create an image. What is the kinetic energy of the electrons after the accelerating process, as they are moving toward the screen?

__________ J

14.3 Electrons accelerate at a high rate from the negative to the positive electrode of an x-ray tube, and emit x-rays when they slam into a metallic target. If the potential difference between the electrodes is 125,000 V, then (a) how much work is done by the field on each electron? (b) What is the kinetic energy of an electron when it reaches the positive electrode?

(a) __________ J
(b) __________ J

14.4 Use the simulation in the interactive problem in this section to calculate the electric potential of plate A needed to accelerate the ion to the desired speed.

__________ V

Section 16 - Equipotential surfaces

16.1 A point charge of 8.40 mC is surrounded by an equipotential surface with a radius of 0.298 m. What is the electric potential on the surface?

__________ V

16.2 What is the radius of the 2.5×10^{-9} V equipotential surface that surrounds an isolated proton?

__________ m

16.3 A positive point charge is surrounded by two equipotential surfaces. (a) What is the shape of the surfaces? (b) On which of the surfaces is the electric potential greater? (c) If the larger of the surfaces has 10 times the area of the smaller, what is the ratio of the potential on the larger to the potential on the smaller?

(a) i. Cube
ii. Pyramid
iii. Sphere
(b) i. The surface closer to the point charge
ii. The surface farther from the point charge
(c) __________

Section 17 - Electric potential and a uniform electric field

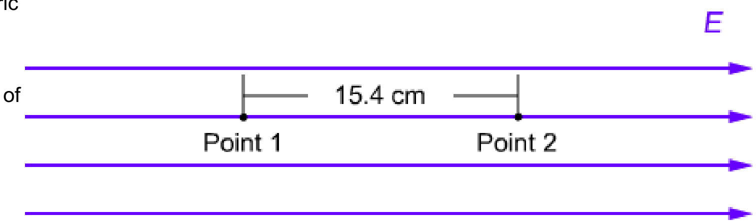
17.1 A uniform electric field of 2.0×10^5 N/C is oriented in the negative *x* direction. How much work must be done to move an electron 3.0 m in the positive *x* direction, if it begins and ends at rest?

__________ J

17.2 In a typical television or in an older computer picture tube (CRT), the electrons that strike the screen to form an image are accelerated between two charged plates that have a potential difference of 2.5×10^4 V. The plates are separated by a distance of 1.3 cm. Assuming that there is a uniform electric field between the plates, calculate the strength of the field.

__________ N/C

17.3 Points 1 and 2 are located in a uniform electric field of strength 325 N/C. The displacement vector from point 1 to point 2 points in the same direction as the field, and has a length of 15.4 cm. The electric potential at point 1 is 9.15 V. What is the potential at point 2?



__________ V

17.4 Starting at a point in a uniform electric field of strength 9.50 N/C, how far and in what direction (relative to the field) must one travel to experience an increase in potential of 13.6 V?

__________ m

i. In the direction of the field
ii. Opposite the direction of the field
iii. Perpendicular to the field

17.5 In the diagram you see the spark gap between the contacts of an automobile spark plug. When the contacts are charged to a potential difference great enough to cause a spark to jump across the gap, the electric field between them approximates a uniform field of strength 3.1×10^6 N/C. What is the magnitude of the potential difference that causes the spark?

__________ V

17.6 An important factor in many biochemical cell processes is the potential difference across a cell membrane. The inside of a cell is at a lower electric potential than the outside, with a potential difference of 0.080 V between them. The membrane consists of two layers of chainlike lipid molecules packed together like the blades of grass in a lawn, but it is only 7.5 nm thick. What is the strength of the electric field inside the membrane?

__________ N/C

17.7 A uniform field has strength 66.8 N/C and points in the positive *x* direction in three-dimensional space. The electric potential at the point (2.50 m, 0 m, 0 m) is −17.3 V. (a) What is the shape of the equipotential surfaces for this field? (b) Write an equation to describe the set of points for which the electric potential is zero volts. (c) Write an equation to describe the set of points for which the electric potential is +24.8 V.

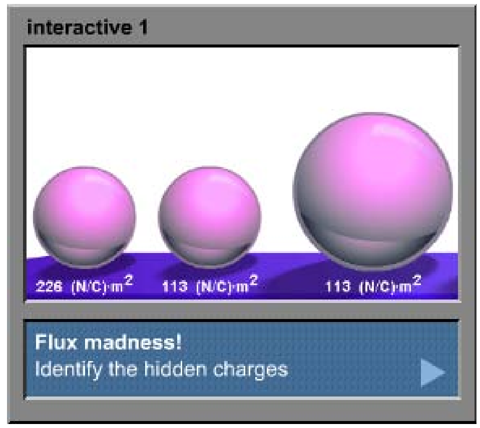
(a)
- ○ Spheres
- ○ Planes parallel to the *xy* plane
- ○ Planes parallel to the *yz* plane
- ○ Planes parallel to the *xz* plane

(b)
- ○ $x = 0$ m
- ○ $z = 0$ m
- ○ $x = 2.24$ m
- ○ $x = 2.50$ m
- ○ $x = 2.78$ m

(c)
- ○ $x = 2.48$ m
- ○ $y = 2.48$ m
- ○ $z = 2.48$ m
- ○ $x = 1.87$ m
- ○ $x = 2.50$ m

Electric Flux and Gauss' Law

26.0 - Introduction



In today's increasingly wired world, more devices than ever create electric fields around us at all hours of the day and night. An important subject in the study of electric fields is the topic of *electric flux*. Electric flux is the product of the electric field strength and the amount of a surface area perpendicular to the field. Electric flux provides an important tool for relating electric charge to electric field.

We will use light striking a white paper card to provide an analogy to electric flux. There are various ways to change the amount of "light flux" at the card's surface. A stronger light illuminates the card more brightly than a weaker one the same distance away. For a given source of light, more light strikes a large card at a given distance than a small card. A card of any size looks brighter if it is turned fully toward the light source, rather than being held at an angle to it.

These properties are all analogous to electric flux. For example, electric flux increases with the strength of the electric field and the amount of surface area through which it passes.

The surface area the light strikes does not have to be flat. If the light source is centered inside a spherical paper lantern, we can ask, "How much light passes through the entire surface of the lantern?" This amount of light is proportional to the brightness of the light bulb it encloses. Perhaps surprisingly, the amount of light passing through the lantern is independent of its surface area. If the lantern is larger, the light is not as intense at the lantern's surface, but the increased surface area means the amount of light passing through the lantern stays the same.

The focus of this chapter is Gauss' law, which makes a statement about the relationship of charge and electric flux that is comparable to the points made in our lantern example. This law states that the total amount of field passing through a closed surface (a surface like a paper lantern or a sphere) is proportional to the amount of charge the surface encloses. The size and shape of the surface do not matter; the amount of electric flux will be the same for any given amount of charge. Gauss' law is an important tool for analyzing the nature of the electric fields generated by a variety of configurations of charge.

You can begin your experimentation with the topic in the simulation here, or you may choose to read a few more sections before using this simulation. The simulation has four spheres, each containing a positive charge. The amount of flux passing through each sphere is shown below it. Two spheres contain one amount of charge, and the other two contain a lesser amount of charge.

In the simulation, you will find two labels displaying the number "+2.00," and two other labels displaying the number "+1.00." You are asked to drag a label to each sphere, indicating which ones you think contain twice as much charge as the others. (In case you are curious, the charge amounts are measured in nanocoulombs.)

Can you deduce which spheres contain the stronger charges? If you can determine this, you already have a strong start toward a fuller understanding of Gauss' law. In any case, read on!

26.1 - Flux

Flux: The amount of field passing through a surface area.


In this chilly pasture, each horse has oriented itself to maximize the flux of winter sunlight.

The concept of flux can be used to describe a relationship between a field and its source. In this chapter, we use flux to relate electric fields to the electric charges that cause them. We begin by discussing what is meant by flux.

Flux is the amount of field passing through a given surface area. The diagram in Concept 1 illustrates this using a uniform field. A uniform field has the same strength and direction at every point. This field is represented using an electric field diagram. Remember, the closer the field lines are to one another, the stronger the field. In other words, the more field lines there are passing through the surface area, the greater the flux.

Flux can be defined for any vector field, including electric and magnetic fields. The symbol for electric flux, Φ_E, is the capital Greek letter phi (pronounced "fie"), with a capital E subscript.

The field shown in Concept 1 is perpendicular to the rectangular surface. The orange color of the field lines is used to indicate that they are passing through the surface. When the field is perpendicular to the surface, the flux equals the product of the field strength and the surface area.

Fields are not always so conveniently oriented however. They may be oriented at an angle to a surface. To account for such situations, we use a more precise definition of flux: It is the amount of field passing **perpendicularly** through a surface area.

We use the concept of an *area vector* to account for the field orientation in determining the flux. You see an area vector in Concept 2. The area vector is perpendicular (normal) to the surface area. Its magnitude equals the amount of surface area.

The angle θ, also shown in Concept 2, is the angle between the field and area vectors. The cosine of θ determines the component of the field passing perpendicularly through the surface. To calculate the flux, use the formulas in Equation 1: The flux equals the dot product of the field and area vectors. To calculate the amount of flux, you multiply the field strength by the surface area and the cosine of θ.

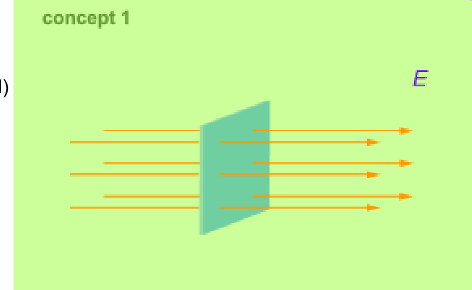
Let's consider this formula in two extreme cases. When the field is parallel to the surface, and perpendicular to the area vector, no field passes through the surface, and the flux equals zero. The equation for flux confirms this: The angle between the field and area vectors is $90°$, and the cosine of $90°$ equals zero. (Note the role played by vectors here: θ is the angle between the field and area vectors.) When the field is perpendicular to the surface, the angle between the vectors is $0°$, the cosine of the angle is one, and the flux is maximized. At this angle the flux equals the product of the field strength and the surface area.

When we say flux, we typically mean "net flux" (just as we should always say that "net force" equals mass times acceleration). Since fields are vectors, two fields will generate fluxes of opposite signs if they pass through the same surface in opposite directions. In this case, the net flux is less than the flux produced by either field alone. (One can also say the flux equals the **net** field passing through a surface.) We will continue to use "flux" in most cases, but will be careful to say "net flux" when fields are combined.

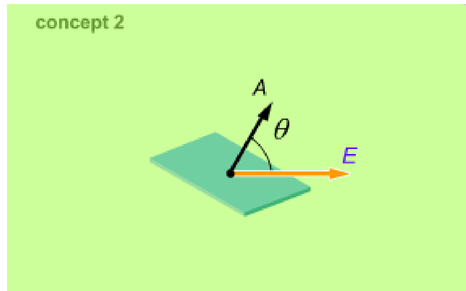
Flux is a scalar quantity. It can be positive or negative, but it does not have a direction. Understanding when it is positive or negative is the topic of another section. Electric flux is measured in newtons per coulomb times meters squared. This is the product of the units for electric field and area. Other kinds of vector fields generate fluxes with different units.

If the concept of flux still confuses you, consider another metaphor. Imagine you are fishing with a net, and a school of fish is swimming at a constant velocity toward your net. The fish, conveniently enough, have a uniform spacing between them. If you have a larger net, one with more surface area, you will catch more fish. If you orient the net so the plane of the net is parallel to the direction in which they are swimming, you will not catch any fish; you must orient it perpendicularly to their direction to maximize your catch. (To use the terminology of the flux equation, you maximize your catch when the net's area vector is parallel to the fishes' velocity vectors.) These ideas are analogous to ones concerning flux: Flux increases with surface area, and is maximized for a given surface when the surface is perpendicular to the field, that is, when the area vector is parallel to the field.

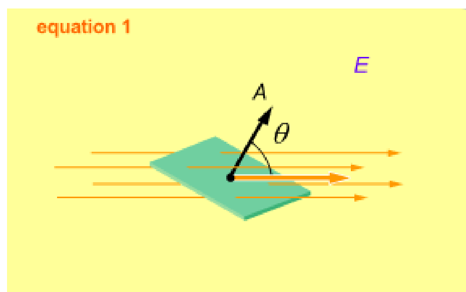
If the fish swim faster, increasing their velocity, more will swim into your net while you are fishing. You can imagine each fish having its own velocity vector. Increasing their velocities is analogous to increasing the field strength: Both actions increase flux. Finally, if all the fish suddenly reverse direction, the fish you caught will escape. This is the equivalent of negative flux.



Flux
Amount of field passing through surface
Depends on:
· Field strength at surface
· Amount of surface area
· Angle between field, area vector



Area vector
Perpendicular to surface
Magnitude equal to area
θ is angle between field, area vector



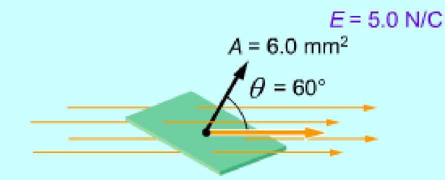
Area vector

$$\Phi_E = \mathbf{E} \cdot \mathbf{A}$$
$$\Phi_E = EA \cos \theta$$

Φ_E = electric flux
$\mathbf{E}$ = electric field
$\mathbf{A}$ = area vector
θ = angle between field, area vector

Units: $\dfrac{\text{newton} \cdot \text{meters}^2}{\text{coulomb}} = \dfrac{\text{N}}{\text{C}} \cdot \text{m}^2$

example 1



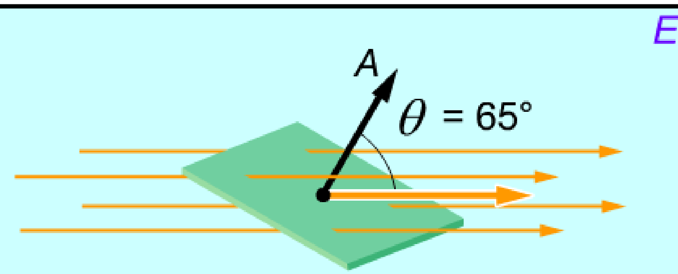
What is the electric flux through the surface?

$\Phi_E = EA\cos\theta$

$\Phi_E = (5.0\,\frac{N}{C})(6.0\times10^{-6}\,m^2)(\cos 60°)$

$\Phi_E = 1.5\times10^{-5}$ (N/C)·m²

26.2 - Interactive checkpoint: electric flux



A uniform electric field fills a region of space. A square piece of paper with sides of length 7.80 cm is held so that its area vector makes a 65.0° angle with the electric field. The flux through the paper is 1.77×10^3 N·m²/C. What is the strength of the electric field?

Answer:

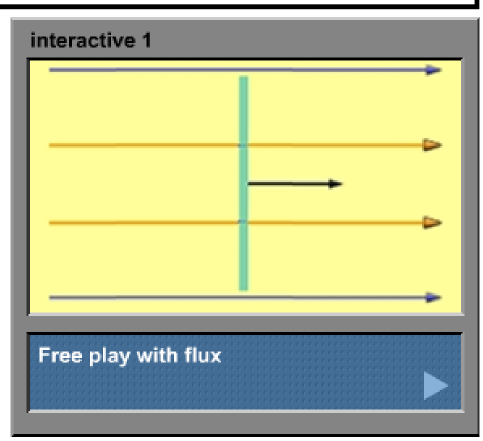
E = ______ N/C

26.3 - Interactive problem: determining flux

To the right are four simulations designed to help you develop and test your understanding of how electric flux is determined by three factors: the strength of the electric field, the amount of surface area, and the angle of intersection between the field and the area vectors.

The first simulation does not ask you to solve a problem, but instead provides an opportunity for free experimentation with the concept of flux. You see a rectangular surface in a uniform electric field. You use the controls in the simulation to change the field strength, the amount of surface area, and the angle of intersection θ between the electric field and surface area vector. A gauge tells you the amount of flux passing through the surface. The orange lines indicate the part of the field that passes through the surface, while the blue lines represent the field that does not pass through the surface.

You can change the angle at which you view the field and surface area with the viewing angle control. You will see that the field lines have arrowheads, which indicate their direction, and tails, represented by ×'s.



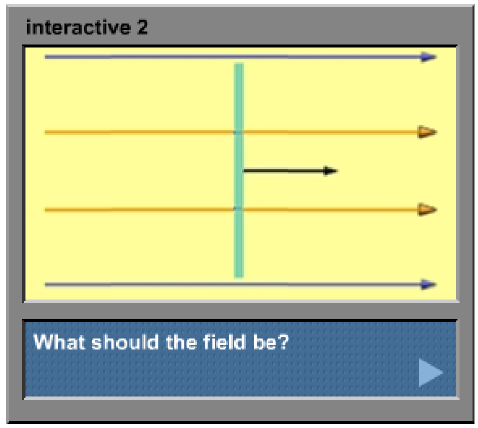
In the next three simulations, labeled as Interactives 2 through 4 on the right, we freeze two of the three controls at certain values and ask you to manipulate the third to produce a flux of **10.0** (N/C)·m².

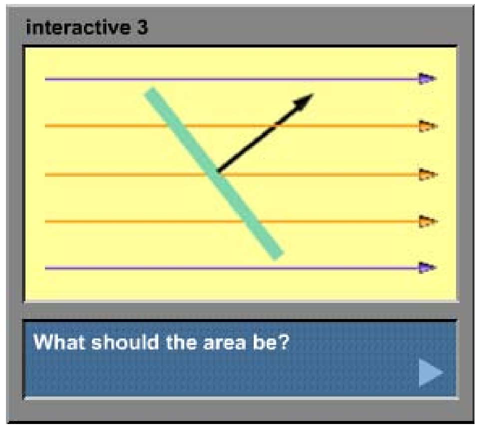
In Interactive 2, we freeze the area of the rectangle at 2.0 m² and the angle at $0.0°$. (This means the field is parallel to the area vector, which means it is perpendicular to the surface.) You job is to set the field strength that will result in the desired flux value. Adjust the field strength control and press CHECK to see the resulting amount of flux displayed on the readout gauge.

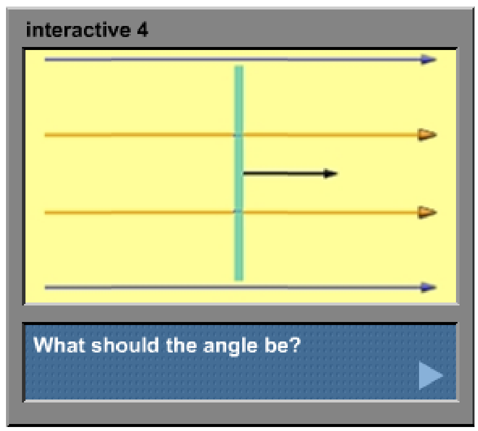
In Interactive 3, the angle θ is $38°$ and the field strength is 4.7 N/C. Set the surface area control to produce a flux of 10.0 (N/C)·m² and again press CHECK to confirm your answer.

In the final simulation, Interactive 4, the area of the rectangle is 2.6 m², and the field strength is 4.7 N/C. Your task here is to set the angle between the field and the area vector to once again create a flux of 10.0 (N/C)·m². As before, set a value and press CHECK.

Interactives 2 through 4 are based on the definition of electric flux, which was stated and discussed in a previous section.

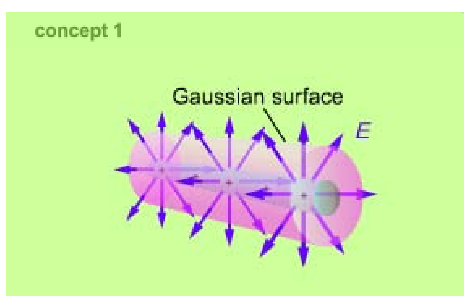






26.4 - Gaussian surfaces

Gaussian surface: A surface that completely encloses a region of space.



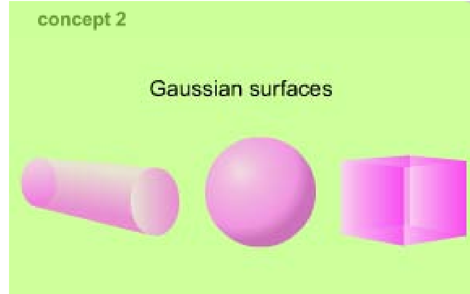
Gaussian surface
Completely encloses a volume of space
Used to measure electric flux

When you think of relating electric charges and electric fields, you likely think of starting with a known configuration of electric charge and calculating the electric field generated by that charge. The German mathematician Carl Friedrich Gauss (1777-1855) took the opposite tack. In the early 19th century, he showed how to analyze electric fields in order to characterize the charges that produce them.

Today, scientists refer to one of his analytical tools as a Gaussian surface. A Gaussian surface is created in the mind of a scientist like you. It can be any shape − sphere, cylinder, box − that completely encloses a volume of space. Such a surface is often referred to as “closed.” If you are unsure whether a surface is closed, imagine filling it with water. If you can rotate the object to any orientation without water leaking out, then the surface is Gaussian. A bowl, for example, would not fit this description.

You can choose a symmetrical shape like a sphere for your Gaussian surface or, if you are feeling more whimsical, you could choose the outside of a football, the exterior of a penguin, or the (unopened) peel of a banana. Any shape that completely surrounds a volume is a Gaussian surface. You see some Gaussian surfaces to the right, with Concept 2 displaying a few typical surfaces, and Concept 3 showing some surfaces less frequently used in applying Gauss' law! Remember: In all cases, the flux is the field passing through the **surface** of these shapes.

Gaussian surfaces are used to analyze the electric fields that pass through them. When a Gaussian surface encloses electric charges, you can make predictions about the nature of those charges by determining the amount of field passing through the surface. Sometimes by choosing a certain shape you can make it easier to study a particular configuration of electric charges. We have not found a case yet where the exterior of a banana or a penguin are particularly useful shapes, but you never know. (If you come up with a use for these shapes, please contact us at gaussian_surfaces@kbooks.com.)



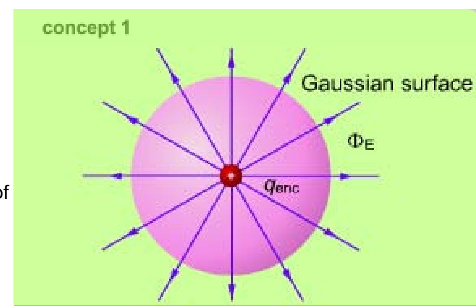
Gaussian surfaces
Can be any shape

Gaussian surfaces
Some shapes more useful than others

26.5 - Gauss' law

Gauss' law: The net electric flux through a closed (Gaussian) surface is proportional to the amount of charge it encloses.

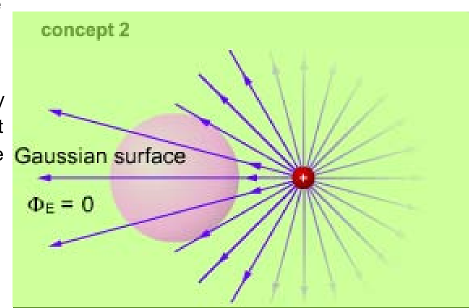
Gauss' law states that the charge enclosed by a Gaussian surface equals the product of the net electric flux passing through the surface and the *permittivity constant*. This is shown in Equation 1. The permittivity constant is written as ε_0 and equals 8.854 19×10^{-12} C^2/N·m^2.



Gauss' law
Flux through Gaussian surface proportional to enclosed charge

In Concept 1, you see a charge enclosed by a Gaussian surface, a sphere in this case. The flux passing through this surface is proportional to the strength of the enclosed charge. The field lines provide a way to "see" Gauss' law at work. The greater the enclosed charge, the stronger the field. A stronger field is represented by closer spaced field lines, which means that more field lines pass through the surface and the flux is greater. This is visual confirmation of Gauss' law: Flux increases with the amount of the enclosed charge.

A spherical Gaussian surface provides a good context for explaining positive and negative flux in greater depth. By convention, the area vector points outward from every element of a Gaussian surface. This means that when a positive charge is positioned at the center of a Gaussian sphere, as in Concept 1, the field and area vectors point in the same direction everywhere on this surface and the angle between them is $0°$. The cosine of this angle is one, so the flux is positive for a positive enclosed charge.



External charge
Creates no net flux

If a negative charge were enclosed, then the field lines would point inward. Since the angle between the field and area vectors would be $180°$ everywhere on the spherical surface, and the cosine of $180°$ is -1, the flux would be negative.

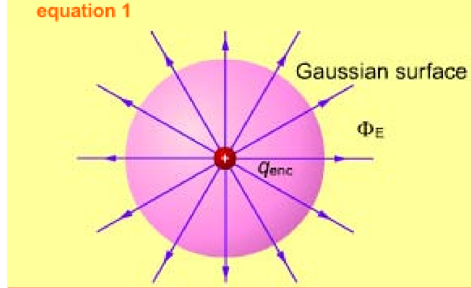
It may seem that altering the shape or size of the Gaussian surface would alter the flux. For example, what would occur if you enclosed the same charge with a larger sphere? Would more flux pass through? The answer is no. You would measure the same amount of flux. The larger sphere would have more surface area, but the field strength would be weaker at its surface. To consider this visually, imagine enlarging the sphere shown in Concept 1. This does not change how many field lines pass through it.

You can perform the same thought experiment mathematically using the case of a spherical Gaussian surface enclosing a positive point charge. Its area increases as a function of the square of the radius (surface area equals $4\pi r^2$), but the field strength decreases at a similar rate (since it equals kq/r^2), so the product of these two values, which equals the flux, is independent of r. We are making this point in a specific case, but it holds true in general that the dimensions of the enclosing shape do not alter the amount of flux that passes through it.

In Concept 2, we show the effect of a charge outside a Gaussian surface. One perhaps surprising implication of Gauss' law is that charges **outside** the surface do not alter the net flux through the surface; only charges inside the surface contribute to the net flux.

At this point you may be thinking that a charge outside the surface would have to cause flux, since field lines from this charge would pass through the surface. You are partially correct: There is flux. However, Gauss' law states that the enclosed charge is proportional to the **net** flux. Consider the field lines from a point charge outside the sphere. Each electric field line that enters the sphere also leaves it. Since as many lines exit as enter, the net flux from the external charge is zero.

equation 1



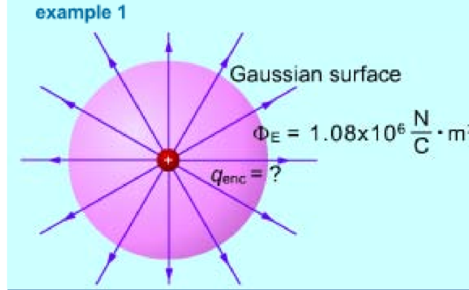
Gauss' law

$$q_{enc} = \varepsilon_0 \Phi_E$$

q_{enc} = enclosed charge
ε_0 = permittivity constant
Φ_E = flux through Gaussian surface

Constant $\varepsilon_0 = 8.85\times10^{-12}\ C^2/N\cdot m^2$

example 1



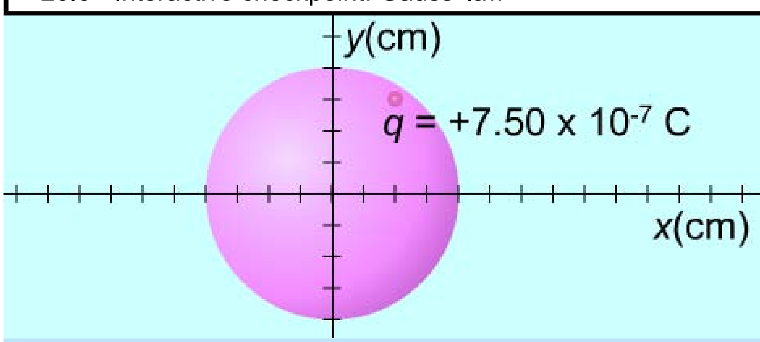
What is the enclosed charge?

$q_{enc} = \varepsilon_0 \Phi_E$

$q_{enc} = (8.85\times10^{-12}\frac{C^2}{N\cdot m^2})(1.08\times10^6\frac{N}{C}\cdot m^2)$

$q_{enc} = 9.56\times10^{-6}\ C = +9.56\ \mu C$

26.6 - Interactive checkpoint: Gauss' law



A sphere of radius 4.00 cm is centered at the origin. There is a point charge of $+7.50\times10^{-7}$ C within the sphere at the point (2.00, 3.00, 0.00) cm. What is the electric flux through the surface of the sphere?

Answer:

Φ_E = ______ $N\cdot m^2/C$

26.7 - Interactive problem: Gaussian spheres

To the right you see three simulations. In the graphic used to launch each simulation, you see a question. You will need to use the simulation to answer the question.

In the first two simulations, you can click on any part of a Gaussian sphere to see the area vector of a surface element at that point, the angle between the area vector and the electric field vector there, and the strength of the electric field. At the risk of being repetitive, the electric field

vector is not shown; the vector shown is the area vector. The strength of the electric field is shown.

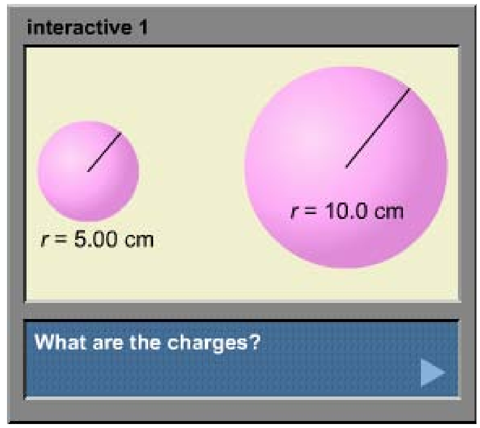
There are two Gaussian spheres in the first simulation, one with a radius of 5.00 cm and the other with a radius of 10.0 cm. Each sphere contains an unknown charge at its center. The spheres are isolated from each other so that each charge's field is kept distinct.

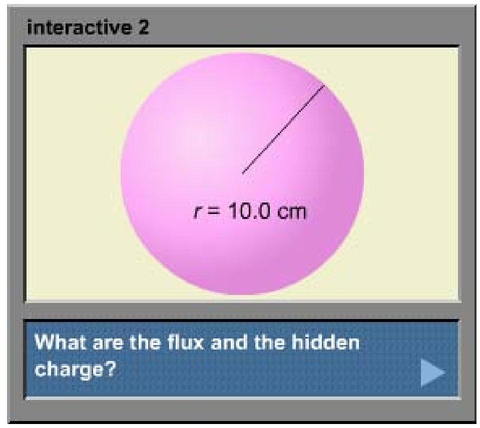
There is a gauge above each sphere that indicates the amount of flux passing through the sphere. You are asked to determine the charge at the center of each Gaussian sphere. Enter the charge in each sphere in the spaces provided in the simulation, to the nearest 0.01 coulomb, and press CHECK. You may find that the flux gauge provides all the information needed to solve the question posed by the simulation.

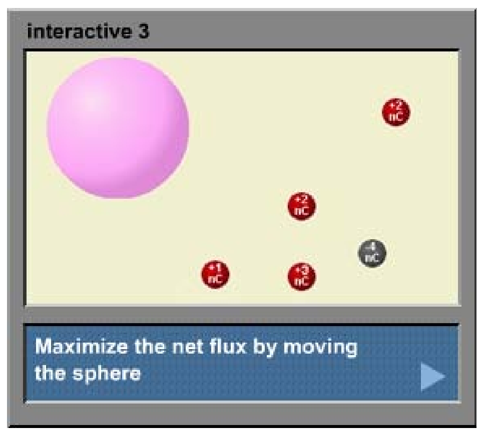
In Interactive 2, there is a sphere with a radius of 10.0 cm, also with a concealed, unknown charge at its center. In this simulation, you are still able to sample the electric field on the surface of the sphere, but there is no gauge to indicate flux. What is the total electric flux through this sphere? What is the magnitude of the charge concealed at its center? Enter these values in the simulation, to the nearest (N/C)·m^2, and to the nearest 0.01 coulomb, and press CHECK.

Interactive 3 has a single Gaussian sphere. You can drag it to various positions on the screen. There is also a collection of charges on the surface of the screen. Your challenge is to position the sphere, enclosing whichever charges you like, so that the flux passing through it is as large as possible. (A charge is enclosed if its center is within the sphere, but you need not go to such a level of precision to solve the problem; you can solve it by enclosing entire charges.) Press CHECK to evaluate the total flux after you have put the sphere where you like. Press REVEAL if you are stumped.

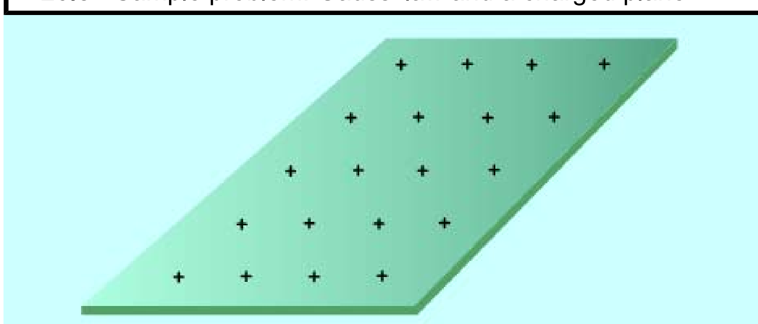
For any of the simulations, if you need help in formulating your answer, return to the section on flux and the section on Gauss' law.







26.8 - Sample problem: Gauss' law and a charged plane



Find the electric field on either side of this plane as a function of the distance from the plane, and of its positive surface charge density σ.

Gauss' law can be used to calculate the strength of the field due to an infinite plane of charge. In the drawing above, you can see a portion of the infinite plane. A positive charge is uniformly distributed across the surface. The positive *surface charge density* − that is, the charge per unit area − is represented by σ. This configuration is often used to model the field created by a finite flat plate, as you will see below.

Draw a field diagram

The field diagram for the charged plane is drawn in the illustration for Equation 1. The field is perpendicular to the surfaces of the plane. The field is uniform: Its strength does not change with distance from the plane. This fact is proved in the derivation below.

Draw a Gaussian surface

Equation 2 shows a Gaussian surface − in this case a rectangular box − that cuts through the plane, enclosing a rectangular section whose area is A. The box is placed so that its top and bottom surfaces are equidistant from and parallel to the plane. The other sides of the box are perpendicular to the plane.

Variables

surface charge density	σ
strength of plane's electric field	E
area of top and bottom surfaces	A
flux through surface	Φ_E
angle of field to area vector	θ
charge enclosed in box	q_{enc}
permittivity constant	$\varepsilon_0 = 8.85\times10^{-12}\ C^2/N\cdot m^2$

What is the strategy?

1. Draw a field diagram. This has already been done in Equation 1.
2. Draw a Gaussian surface. This has been done in the subsequent illustrations to the right. Observe where the field is parallel or perpendicular to the Gaussian surface, and where it is constant on the surface.
3. Calculate the flux through the Gaussian surface. The symbol E for the field strength will remain in your expression for flux.
4. Apply Gauss' law to relate the flux to the charge enclosed within the Gaussian surface. Solve for E to derive an equation for the strength of the field.

Physics principles and equations

We will use the definition of electric flux through a surface,

$$\Phi_E = EA\cos\theta$$

We also use Gauss' law,

$$q_{enc} = \varepsilon_0 \Phi_E$$

The definition of surface charge density is charge per unit area:

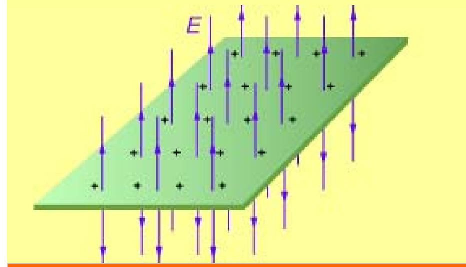
$$\sigma = q/A$$

Step-by-step solution

We have drawn the electric field. It is perpendicular to the plane, uniform, and directed away from it on either side. The Gaussian surface is a box bisected by the plane. The field is parallel to the area vectors of the top and bottom surfaces of this box, and its strength is everywhere the same on them. The field is perpendicular to the area vectors of the box's side surfaces.

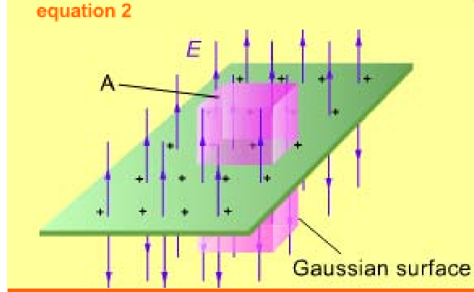
Step	Reason
1. **E** parallel to area vectors of top and bottom surfaces	by inspection
2. **E** perpendicular to area vectors of side surfaces	by inspection
3. E the same on top and bottom surfaces	by inspection and symmetry

equation 1



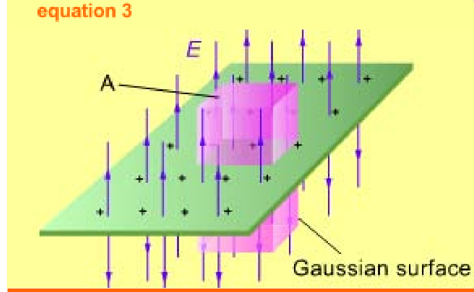
The field diagram

equation 2



The Gaussian surface and the flux

equation 3



Electric field of an infinite plane

$$E = \frac{\sigma}{2\varepsilon_0}$$

E = electric field strength
σ = surface charge density of plane
ε_0 = permittivity constant

In the following steps we calculate the flux Φ_E through the Gaussian box.

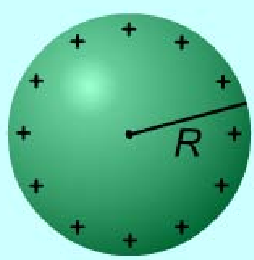
Step	Reason
4. $\Phi_E = EA\cos\theta$	definition of electric flux
5. $\theta \equiv 0°$ on top and bottom	justified by step 1
6. $\Phi_{E,top} = EA(1)$	use angle from step 5 in equation 4
7. $\Phi_{E,bottom} = EA(1)$	use angle from step 5 in equation 4
8. $\theta \equiv 90°$ on sides	by inspection
9. $\Phi_{E,total} = 2EA$	no flux through sides of box

The expression for $\Phi_{E,total}$ derived in the previous step contains the field strength E. We apply Gauss' law to relate the flux to the enclosed charge, then solve for E, to determine the strength of the field.

Step	Reason
10. $q_{enc} = \sigma A$	enclosed charge is surface density times area
11. $q_{enc} = \varepsilon_0 \Phi_E$	Gauss' law
12. $\sigma A = \varepsilon_0(2EA)$	substitute equations 10 and 9 into equation 11
13. $E = \dfrac{\sigma}{2\varepsilon_0}$	solve for E

The equation shows that for an infinite charged plane, the field strength does not change with distance from the plane. This enables us to model the field of a flat charged plate as a uniform field (one that does not change with position − in this case, distance from the plate). This proves useful in analyzing devices such as capacitors, which in their simplest form are modeled as two oppositely charged plates separated by a small distance. Treating the field of a finite flat plate as if it were uniform is a good approximation for distances from the plate much smaller than its dimensions, and for points not close to its edges.

26.9 - Sample problem: Gauss' law and a charged sphere



Describe the electric field outside and inside a charged conducting sphere as a function of the distance from the center of a sphere. The charge q is uniformly distributed on the surface of the sphere.

In this section, we will use Gauss' law to find the strength of the electric field generated both outside and inside a charged conducting sphere. We will exploit the similarities between these two problems so that we do not repeat many steps in the solution.

Draw a field diagram

The field diagram for the charged conducting sphere is drawn in the illustration for Equation 1.

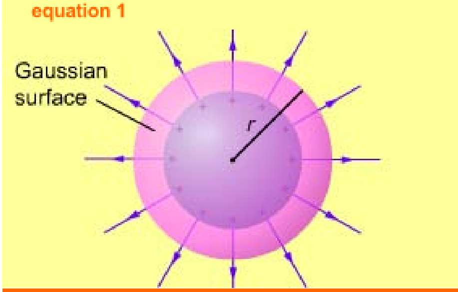
The charge on a conducting sphere will be in electrostatic equilibrium, meaning that the field outside the sphere radiates directly outward in every direction. Furthermore, the charge is uniformly distributed on the sphere's surface, which means the strength of the outside field depends only on the distance from the center of the sphere.

As we will show shortly, there is no field inside the sphere.

Draw a Gaussian surface

In Equation 1, you see a spherical Gaussian surface. This surface has radius r, and it encloses the charged conducting sphere. The electric field is perpendicular to the Gaussian surface, and the fixed radius of the surface means the field strength is constant on it.

equation 1



Outside a positively charged sphere

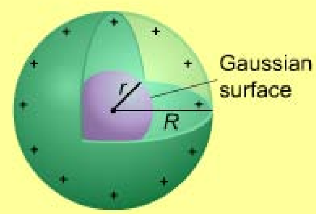
$$E = \frac{q}{4\pi\varepsilon_0 r^2} = \frac{kq}{r^2}$$

Variables

total charge on conducting sphere	q
radius of conducting sphere	R
strength of electric field	E
radius of Gaussian surface	r
flux through Gaussian surface	Φ_E
area of Gaussian surface	A
angle of E-field to area vector	θ
permittivity constant	$\varepsilon_0 = 8.85\times10^{-12}\ C^2/N{\cdot}m^2$
Coulomb's constant	$k = 8.99\times10^9\ N{\cdot}m^2/C^2$

E = electric field strength
q = positive charge on sphere
ε_0 = permittivity constant
r = distance from center of sphere
k = Coulomb's constant

equation 2



Inside conducting sphere

$E \equiv 0$

E = electric field strength

What is the strategy?

1. Draw a field diagram. This has already been done in Equation 1.
2. Draw a Gaussian surface. This has been done, as well. Observe where the field is parallel or perpendicular to the Gaussian surface, and where it is constant on the surface. As stated in the discussion above, the field is perpendicular to the entire surface, and constant everywhere on it.
3. Calculate the flux through the Gaussian surface. Use the results of the previous observations to simplify the calculation as much as possible. The symbol E for the field strength will remain in your expression for flux.
4. Apply Gauss' law to relate the flux to the charge enclosed inside the Gaussian surface. Solve for E to obtain an equation for field strength.

Physics principles and equations

Excess charge is uniformly distributed on the surface of a conductor in electrostatic equilibrium, and the external field generated by the charge extends radially outward from the surface of the conductor.

We will use the definition of electric flux through a surface,

$$\Phi_E = EA\cos\theta$$

θ is zero everywhere on the surface outside the conducting sphere.

We also use Gauss' law,

$$q_{enc} = \varepsilon_0\Phi_E$$

The formula for the area of a spherical surface is

$$A = 4\pi r^2$$

Coulomb's constant and the electrostatic permittivity constant are related by the equation

$$k = 1/4\pi\varepsilon_0$$

Step-by-step solution − field outside the sphere

In the following steps we calculate the flux Φ_E through the spherical Gaussian surface.

Step	Reason
1. $\Phi_E = EA\cos\theta$	definition of electric flux
2. $\theta \equiv 0°$ on surface	field outward, perpendicular to sphere
3. $\Phi_E = E(4\pi r^2)(1)$	substitute area of sphere and constant angle into equation 1

The expression for Φ_E derived in the previous step contains the field strength E. We apply Gauss' law to relate the flux to the enclosed charge, then solve for E.

Step	Reason
4. $q_{enc} = \varepsilon_0 \Phi_E$	Gauss' law
5. $q = \varepsilon_0(4\pi r^2 E)$	substitute charge q and equation 3 into equation 4
6. $E = \dfrac{q}{4\pi\varepsilon_0 r^2}$	solve for E
7. $E = \dfrac{kq}{r^2}$	substitute Coulomb's constant

Note an interesting feature of the solution: The field strength is the same as it would be for a positive point charge q at the center of the sphere. This confirms a form of the *shell theorem*: When calculating an electrostatic force or field outside a charged sphere, the sphere's charge can be treated as though all of it resides at the center point of the sphere.

As with a point charge, the magnitude of the electric field outside a positively **or negatively** charged conducting sphere can be expressed as

$$E = \frac{k|q|}{r^2}$$

When we describe the field outside a charged conducting sphere, it does not matter whether the sphere is a shell or a solid. In either situation, the charge is uniformly distributed on its surface, since charge always accumulates uniformly on the surface of a conducting object.

We now turn our attention to the second problem posed above: What is the field **inside** a charged conducting sphere? We use the illustration for Equation 2 to help answer this question.

Step-by-step solution − field inside the sphere

The expression for Φ_E contains the field strength E. We apply Gauss' law to relate the flux to the enclosed charge (zero), then solve for E, to complete the description of the field by stating its value inside the charged sphere.

Step	Reason
1. $q_{enc} = \varepsilon_0 \Phi_E$	Gauss' law
2. $0 = \varepsilon_0 EA \cos\theta$	zero enclosed charge and definition of flux
3. $E = 0$	solve for E

The field strength is zero everywhere inside a charged conducting sphere.

26.10 - Electric field of two infinite planes: a capacitor

A device called a parallel-plate capacitor can be modeled as two parallel infinite planes. The planes have opposite charges but the magnitudes of their charge densities are the same. This configuration, shown in the diagram to the right, is called an *ideal parallel-plate capacitor*. (We magnify a section of such a capacitor for the sake of visual clarity; the plates are quite close together in typical capacitors.)

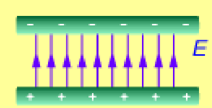
Fields obey the principle of superposition. That is, we can add them as vectors at any point to determine the net field at that point. We use that principle here to find the field strength between, and outside, the capacitor plates.

An infinite charged plane generates a uniform field perpendicular to it on either side, directed away from a positively charged plane and toward a negatively charged plane. Between a pair of oppositely charged capacitor plates, the fields from both plates point in the same direction, so the combined field between the plates is stronger than either field alone. The field between the plates equals the sum of the fields of each plate. You see a formula for the interior field strength in Equation 1 to the right. It equals two times the uniform field strength due to an infinite plane of charge.

On either side outside the capacitor plates, the fields point in opposite directions. When added, they sum to zero. Since the two fields nullify each other there, an ideal capacitor produces no net field outside the region between the plates.

Again, this analysis of an ideal parallel-plate capacitor is based on the assumption of **infinite** planes. However, when a distance that is small relative to the length and width of two finite plates separates them, the equation shown to the right provides a good approximation for points between the plates that are not near their edges.

equation 1



Parallel-plate capacitor

Field between plates of ideal capacitor

$$E = \frac{\sigma}{\varepsilon_0}$$

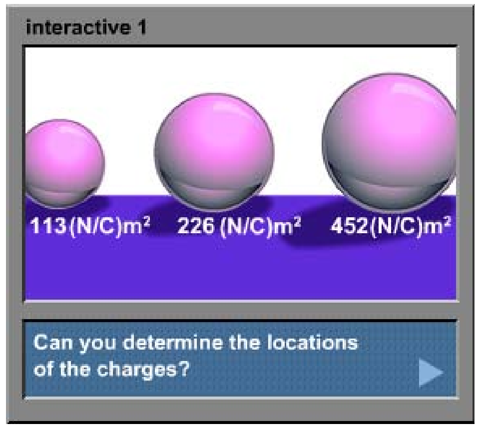
E = field strength between plates
σ = charge density of positive plate
ε_0 = permittivity constant

26.11 - Interactive group problem: shell game

At the right, you see a physicist's version of the old shell game. There are three spheres; they have radii of 4.00, 6.00, and 8.00 cm respectively. There are also charges of positive 1.00, 2.00, and 4.00 nC (nanocoulombs). One player hides the charges and the other has to use the flux gauges to determine the charges contained by each sphere. If you determine which shell contains which charges, you win!

Play this game against a friend. He or she can drag the charges in any combination to hide them in the spheres. The charges disappear when they are inside the spheres.

You can then use Gauss' law, introduced in an earlier section, together with the flux readings to determine how many charges are in each sphere. A shell can contain none of the charges, three of them, or any number in between. You drag the labels provided to show where you think the charges are hidden. Happy hunting! And can we interest you in a little wager?



26.12 - Gotchas

Confusing flux and field. Flux measures how much field is passing through a given surface area. This is reflected in the units for flux: (newtons per coulomb) × (meters squared). Newtons per coulomb are the units for electric field, and meters squared are the units for area.

If a field intersects a surface perpendicularly, the angle of intersection is $90°$. Yes, but... when physicists speak of the angle θ for the purpose of calculating flux, they mean the angle between the field and an **area vector** normal to the surface. In the configuration described, that angle θ is $0°$. If the field is parallel to the surface, then $\theta = 90°$.

A paper cup is a Gaussian surface. No, since it does not completely enclose a volume. A Gaussian surface is a "closed surface" that encloses a volume. To judge whether a given surface is Gaussian or not, visualize filling it with water. If it cannot leak, no matter how you rotate it, then it is a Gaussian surface. Water will spill out of a cup.

A charged particle outside a Gaussian surface generates no flux through the surface. You should say it generates no **net** flux through the surface, since the charge does cause field to pass through the surface, resulting in positive and negative flux that sums to zero. But if you said "yes", you likely meant the right thing.

26.13 - Summary

Flux is a measure of the amount of field that passes through an imaginary surface. It can apply to any kind of field, and in general is symbolized by the capital Greek letter phi, Φ. It is a scalar.

The area vector corresponding to a flat surface or surface element is perpendicular to the surface, and has a magnitude equal to the area of the surface.

Electric flux, Φ_E, is equal to the dot product of the electric field and the area vector of the flat surface (or surface element) through which the field passes. The units for electric flux are (N/C)·m^2.

Gauss' law states that the net electric flux through a closed surface is proportional to the amount of charge it encloses. Such a surface is called a Gaussian surface. Charges outside the surface do not contribute to the **net** flux through the surface.

Here are some tricks for applying Gauss' law to make your calculations easier: Choose a Gaussian surface such that the field strength is constant over large portions of the surface. Usually this means that those parts of the surface are at a constant distance from the source of the field. Also, the **area vector** always points out of each surface element perpendicularly, so make sure the field makes a constant angle (usually zero) with the area vector over large portions of your surface.

Cylinders, boxes, and spheres are the most useful shapes for Gaussian surfaces. Examples of the usefulness of these shapes are to be found in the analysis of the electrical fields due to an infinite plane of charge or a charged conducting sphere

Two oppositely charged plates placed parallel to each other form a capacitor. The superposition of their fields means that the field between them is twice the field of a single charged plate, and that there is no field anywhere other than in the region between the plates.

Equations

Electric flux

$$\Phi_E = \mathbf{E} \cdot \mathbf{A}$$

$$\Phi_E = EA \cos \theta$$

Gauss' law

$$q_{enc} = \varepsilon_0 \Phi_E$$

Electric field of an infinite plane

$$E = \frac{\sigma}{2\varepsilon_0}$$

Charged conducting sphere: outside

$$E = \frac{k|q|}{r^2}$$

Charged conducting sphere: inside

$$E \equiv 0$$

Between plates of ideal capacitor

$$E = \frac{\sigma}{\varepsilon_0}$$

Chapter 26 Problems

Conceptual Problems

C.1 A thick book is suspended in a uniform nonzero electric field. The field is parallel to the plane of the book's front cover, and the flux through the spine (which has less area than the cover) is positive. Which is greater: the flux through the spine, or the flux through the front cover?

i. The flux through the spine
ii. The flux through the cover
iii. They are equal

C.2 Suppose a positive charge is at the center of a cube. Another positive charge is outside the cube, on a line perpendicular to a face of a cube, that passes through its center. Which face of the cube has the greatest (positive) net flux through it?

- ○ The face farthest from the external charge
- ○ The face closest to the external charge

C.3 Which of the following are Gaussian surfaces?

- ☐ An infinitely thin flat sheet
- ☐ An open tube with infinitely thin walls
- ☐ A capped bottle of root beer with infinitely thin walls

C.4 Which of the following could be used for a Gaussian surface? (Check all that apply.)

The surface of a
- ☐ Sphere
- ☐ Thin cup
- ☐ Donut (torus)
- ☐ Square
- ☐ Bowling pin
- ☐ Ellipse
- ☐ Infinite plane
- ☐ Banana
- ☐ Football

C.5 What is the shape of the Gaussian surface consisting of the points in three dimensional space for which $x^2 + y^2 + z^2 = r^2$?

i. A square
ii. A cube
iii. A cylinder
iv. A cone
v. A sphere
vi. An ellipsoid

C.6 Two cats are initially electrically neutral before they rub against each other in greeting. As a result of the rubbing, one cat acquires a small positive charge, and the other one acquires an equal negative charge. The two cats then explore the inside of a small box, one at a time. What is the relation of the electric flux through the box surface when the positive cat is inside, compared to when the negative cat is inside?

i. They are equal
ii. They have equal magnitudes but opposite signs
iii. They have different magnitudes

C.7 A positive charge is at the center of a spherical Gaussian surface. The strength of the field due to the charge is the same everywhere on the surface of the sphere, and the flux through the sphere equals this strength times the area of the sphere. Now the charge is moved off center, very close to a point on the sphere's surface, but still inside the sphere. Due to the proximity of the charge, the flux through the nearby surface of the sphere increases significantly. Does the total flux through the sphere increase, decrease, or stay the same? Explain your answer.

The flux
i. increases
ii. decreases
iii. stays the same

C.8 Lithium is a light metallic element whose ions are used in highly efficient batteries for portable laptop computers. Each lithium ion (Li^+) contains two electrons and three protons. Suppose a lithium ion is surrounded by a Gaussian surface. Is the flux through the surface positive, negative, or zero?

i. Positive
ii. Negative
iii. Zero

Section Problems

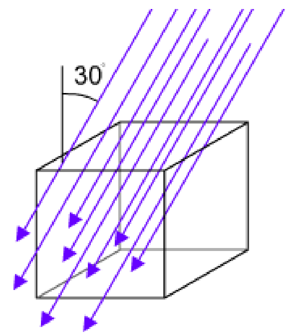
Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Does the flux increase, decrease, or stay the same as the size of the sphere increases, given a constant enclosed charge? (b) Does the flux increase, decrease, or stay the same as the amount of charge increases, given a constant size sphere?

(a) i. Increases
ii. Decreases
iii. Stays the same
(b) i. Increases
ii. Decreases
iii. Stays the same

Section 1 - Flux

1.1 A cube with 1 m edges is in a uniform electric field of magnitude 17 N/C. The electric field intersects the top of the cube at a 30° angle to the normal, and is parallel to two faces of the cube. The area vector for each face of the cube points outward. (a) What is the relationship between the electric flux through the top face and the electric flux through the bottom face? (b) Calculate the net electric flux through the entire surface of the cube.



(a) i. They are equal
ii. They have equal magnitude but opposite signs
iii. They have unequal magnitudes
(b) ________ (N/C) · m^2

1.2 (a) A uniform electric field of strength E = 7.96 N/C passes through a flat surface at an angle of 24.5° (that is, the angle it makes with the area vector is 24.5°). The electric flux is 6.15 (N/C)·m^2. What is the area of the surface? (b) A uniform electric field of strength E = 334 N/C passes through a flat surface at an angle of 0.887 radians. The electric flux is 455 (N/C)·m^2. What is the area of the surface?

(a) ________ m^2
(b) ________ m^2

1.3 (a) An electric field passes perpendicularly through a surface (that is, the field is parallel to the area vector of the surface). The surface is a circle with a diameter of 0.985 m. The flux is Φ_E = 10.7 (N/C)·m^2. What is the strength of the field? (b) An electric field passes at a 45.0° angle through a surface that is an equilateral triangle with an edge length of 1.00 m. The flux is Φ_E = 0.0552 (N/C)·m^2. What is the strength of the field?

(a) ________ N/C
(b) ________ N/C

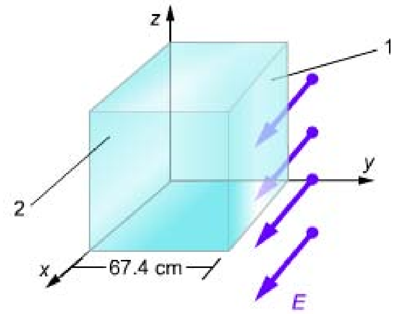
1.4 The area of the loop formed by a wire coat hanger is 0.232 m^2. The coat hanger is immersed in a uniform electric field of 1120 N/C and the flux through the coat hanger is 100 (N/C)·m^2. Determine the angle between the the electric field and the area vector of the plane of the wire.

________ °

1.5 The electric field **E** = (5.4, 2.7, 8.1) N/C passes through a square surface whose area vector is **A** = (1.7, −3.4, 0.55) m^2. What is the electric flux Φ_E through the surface?

________ (N/C) · m^2

1.6 A cube is situated in three-dimensional space as shown, with its edges oriented parallel to the coordinate axes. The length of each edge is 67.4 cm, and the area vector of each face points outward from the cube. The region of the cube is occupied by a uniform field having strength 487 N/C and pointing in the direction of the positive *x* axis. (a) What is the flux through the "upstream" face of the cube, labeled "1" in the diagram? (b) What is the flux through the "downstream" face of the cube, labeled "2" in the diagram? (c) What is the flux through any of the other four faces of the cube? (d) What is the total flux through all the faces of the cube?



(a) ________ (N/C) · m^2
(b) ________ (N/C) · m^2
(c) ________ (N/C) · m^2
(d) ________ (N/C) · m^2

1.7 (a) A cube has six square faces and twelve edges. The length of each edge is 12.3 cm. Each of the six faces has an outward-pointing area vector. What is the sum of the six area vectors? (b) A triangular prism has five faces. The two parallel faces are equilateral triangles, and the other three faces are squares. This solid has nine edges, and the length of each edge is 33.2 cm. Each of the five faces has an outward-pointing area vector. What is the sum of the five area vectors?

(a) ________ m^2
(b) ________ m^2

1.8 Light can deliver energy to a surface that it illuminates, as you can appreciate if you have ever warmed yourself in the Sun on a cold day. The equation for the power of sunlight on a surface is $P = IA \cos \theta$, where P is the power, I is the "intensity" of the sunlight (roughly, its brightness), A is the area of the surface, and θ is the angle between the light and the area vector. This is a "light" analogy of the electric flux through a surface, with power playing the role of flux. (a) The circular disk of a sunflower head has a radius of 13.5 cm. The intensity of direct sunlight at the Earth's surface is 1240 W/m^2. What is the light power P received by the sunflower disk when it points directly at the Sun, that is, when its area vector is parallel to the direction of the sunlight? (b) The photograph of sunflowers that accompanies this problem was taken at 5:00 P.M., when the angle of the Sun from the zenith was 73.3°. If sunflower blossoms always pointed straight up (as daisies do), what power would the sunflower disk described in part "a" be receiving at 5:00 P.M.?

(a) ________ W
(b) ________ W

1.9 If an imaginary surface is immersed in a flow of water (inside a pipe, for example) the volume flux of the water through the surface is defined as the number of cubic meters of water flowing through the surface per second. In a particular water main, the inside diameter of the pipe is 8.24 cm, and water flows through it at a velocity of 3.45 m/s. (a) Use an imaginary surface consisting of a flat disk, completely filling the pipe, with its area vector parallel to the water velocity vector. What is the volume flux of water through the pipe? (b) Repeat the calculation, but rotate the disk, keeping it the same size, and completely within the pipe, so that its area vector makes an angle of 45° with the water velocity vector.

(a) ________ m^3/s
(b) ________ m^3/s

Section 3 - Interactive problem: determining flux

3.1 Use the simulation in the first interactive problem in this section to answer the following questions. (a) If you double the surface area, and keep everything else constant, what happens to the flux? (b) How is the flux related to the surface area? (c) If you double the electric field strength, and keep everything else constant, what happens to the flux? (d) How is the flux related to the electric field strength?

(a) The flux
i. halves .
ii. doubles
iii. quadruples
iv. stays the same

(b) Flux is
i. inversely proportional to surface area.
ii. directly proportional
iii. not related

(c) The flux
i. halves .
ii. doubles
iii. quadruples
iv. stays the same

(d) Flux is
i. inversely proportional to electric field strength.
ii. directly proportional
iii. not related

3.2 Using the information given in the second interactive problem in this section, what field strength will produce the desired flux? Test your answer using the simulation.

__________ N/C

3.3 Using the information given in the third interactive problem in this section, what is the area required to create the desired flux? Test your answer using the simulation.

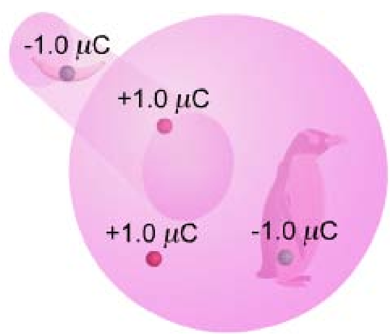
__________ m^2

3.4 Using the information given in the fourth interactive problem in this section, what is the angle required to create the desired flux? Test your answer using the simulation.

__________ °

Section 5 - Gauss' law

5.1 Four Gaussian surfaces are shown, along with four charges. The banana is inside the cylinder, the penguin is inside the sphere, and one end of the cylinder is inside the sphere. Determine the electric flux through the surface of (a) the banana (b) the cylinder, (c) the sphere, and (d) the penguin.



(a) __________ $(N/C) \cdot m^2$
(b) __________ $(N/C) \cdot m^2$
(c) __________ $(N/C) \cdot m^2$
(d) __________ $(N/C) \cdot m^2$

5.2 A cherry with charge 0.00000360 C is embedded in an uncharged cube of jello. Find the flux going through the walls of the jello.

__________ $(N/C) \cdot m^2$

5.3 An uncharged spherical shell has a diameter of 9.42 m, and contains some distribution of charge inside. There is an electric field that points radially outward everywhere on the sphere's surface, with a strength of exactly 244 N/C at all points on the surface. (a) Use Gauss' law to find the total charge enclosed by the spherical shell. (b) Which descriptions of the distribution of charge inside the shell **could** be true? (Check all that apply.)

(a) __________ C

(b)
- ☐ There is a single point charge
- ☐ There are eight point charges at the vertices of a cube
- ☐ The charge is spherically symmetric
- ☐ The charge is uniformly distributed on a line through the sphere's center

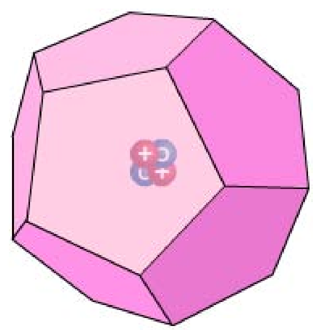
5.4 A charge of +5.85 μC resides at the origin. (a) What is the flux through any Gaussian surface enclosing this charge? (b) What is the strength of the field due to the charge at a distance of 8.00 cm from the origin? (c) The charge is surrounded by a sphere, centered at the origin, of radius 8.00 cm. What is the product of the field strength at 8.00 cm and the area of this sphere? (d) Explain why your answers to parts a and c are the same. (If they are not, check your math!)

(a) ________ (N/C) · m^2
(b) ________ N/C
(c) ________ (N/C) · m^2
(d) Submit answer on paper

5.5 You are blowing bubbles, using a bubble wand that has a circular loop at one end with a radius of 1.40 cm. You hold the wand so that the area vector of the loop points in the same direction as a uniform electric field of 23.2 N/C. (a) What is the electric flux through the flat soap film that spans the loop before you begin to blow? (b) What would the electric flux be if the area vector had pointed in the opposite direction? (c) Now you blow in the same direction as the electric field, and the film bulges out to form the greater part of a spherical surface of radius 3.34 cm. What is the flux through the film now? (d) You blow some more, and the film detaches to form a complete spherical bubble of radius 3.46 cm, which floats off the wand. What is the net flux through the bubble?

(a) ________ (N/C) · m^2
(b) ________ (N/C) · m^2
(c) ________ (N/C) · m^2
(d) ________ (N/C) · m^2

5.6 Your friend encases an alpha particle (charge: +2*e*) exactly in the middle of a dodecahedron with an edge length of 5.0 cm. A dodecahedron is a three-dimensional object that has twelve regular pentagons as its faces, each one having exactly the same area. She challenges you to find the electric flux through one face of this polyhedron. What is your reply? (Hint: First calculate the total flux through the surface and then invoke a symmetry argument.)



________ (N/C) · m^2

Section 7 - Interactive problem: Gaussian spheres

7.1 Use the information given in the first interactive problem in this section to calculate the charge in (a) the left sphere and (b) the right sphere. Test your answer using the simulation.

(a) ________ C
(b) ________ C

7.2 Use the information given in the second interactive problem in this section to calculate (a) the flux and (b) the hidden charge. Test your answer using the simulation.

(a) ________ (N/C) · m^2
(b) ________ C

7.3 Use the information given in the third interactive problem in this section to determine which of the following charge configurations should be enclosed by the Gaussian sphere in order to maximize the flux. Test your answer using the simulation.

i. +2 nC; +3 nC; -4 nC
ii. +1 nC; -1 nC
iii. +4 nC
iv. +2 nC
v. +2 nC; +3 nC
vi. +3 nC; -4 nC

Section 8 - Sample problem: Gauss' law and a charged plane

8.1 An astronaut travels to an alternate universe that contains an infinite charged plane and no other charged objects. The astronaut observes that the electric field in this universe has a magnitude of 2.20 N/C everywhere and points directly away from the plane. Determine the charge density of the plane.

________ C/m^2

8.2 An infinite plane is uniformly charged with static electricity and generates an electric field. The field has a strength of 7.50 N/C at a distance of 1.00 m from the plane. (a) What is the strength of the field 2.00 m from the plane? (b) What is the strength of the field 10.0 m from the plane?

(a) __________ N/C
(b) __________ N/C

8.3 A proton is shot perpendicularly at an infinite plane of charge. The charge density of the plane is $+7.65\times10^{-4}$ C/m^2. If the proton starts out 1.05 mm from the plane, what must its velocity be so that it screeches to a halt exactly as it reaches the plane?

__________ m/s

8.4 An alpha particle with a charge of 3.20×10^{-19} C and a mass of 6.68×10^{-27} kg rests just above a horizontal plane of vast extent having a charge density of 6.21×10^{-9} C/m^2. The electric field generated by the plane will accelerate the particle upward. How high above the plane will the particle be after 2.15×10^{-3} s?

__________ m

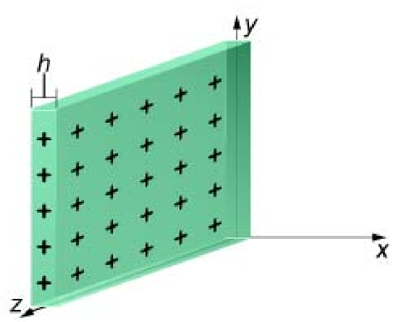
8.5 In this problem you will look at an infinite charged plane from another point of view. Suppose you are in a magical vehicle that hovers above a positively charged plane, and that there is a very small window in the vehicle through which you can view the plane. At a height of 2.00 km, the window affords you a view of a region of the plane 12.4 m^2 in extent. The positive charge density of the plane is small, coming from scattered protons, and its magnitude is 750e/m^2. (a) How many protons are in your view of the plane? (b) Since the angular size of your view is small, you can assume that the electric field vector at your location due to each proton points in the same direction, and you can calculate the net electric field strength due to all the visible protons by simply adding the individual contributions. What is the total field strength due to the protons visible in the window? (c) Now your vehicle zooms up to a distance of 4.00 km above the plane. What is the new area you can see? (d) How many protons are in your enlarged view? (e) What is the total field strength due to the protons visible in the window?

(a) __________ protons
(b) __________ N/C
(c) __________ m^2
(d) __________ protons
(e) __________ N/C

8.6 A vertical, infinite, nonconducting plane has uniform positive charge. A weightless thread is attached to the plane, and at the end of the thread there is a small spherical ball with a charge of +78.6 μC, and a mass of 0.954 kg. The angle that the thread makes with the plane is 37.3°. What is the charge density σ of the plane?

__________ C/m^2

8.7 Consider a rectangular slab, infinite in length and width and extending a distance h/2 on either side of the yz plane, so that it has a finite thickness h. The slab has a constant positive charge density ρ. (a) In what direction does the electric field point for positive x? (b) In what direction does the electric field point for negative x? (c) Determine the electric field inside the slab, for $0 < x < h/2$, using Gauss' Law. (d) Determine the electric field outside the slab, for $x > h/2$, using Gauss' Law.



(a) In the
i. positive x direction.
ii. negative x
iii. positive y
iv. negative y
v. positive z
vi. negative z

(b) In the
i. positive x direction.
ii. negative x
iii. positive y
iv. negative y
v. positive z
vi. negative z

(c) ○ $2\rho x/\varepsilon_0$ ○ $\rho x/\varepsilon_0$ ○ $\rho x/2\varepsilon_0$
(d) ○ $2\rho h/\varepsilon_0$ ○ $\rho h/\varepsilon_0$ ○ $\rho h/2\varepsilon_0$

Section 9 - Sample problem: Gauss' law and a charged sphere

9.1 A spherical shell with a radius of 0.500 m is uniformly charged and generates an electric field. The field has a strength of 7.50 N/C at a distance of 1.00 m from the center of the sphere. (a) What is the strength of the field 2.00 m from the center of the sphere? (b) What is the strength of the field 10.0 m from the center of the sphere? (c) What is the strength of the field inside the shell?

(a) __________ N/C (b) __________ N/C (c) __________ N/C

9.2 (a) A hollow, conducting spherical shell, of radius 3.50 cm, has charge -9.14×10^{-5} C. What is the strength of the electric field 2.50 m from the center of the sphere? (b) What is the direction of the field? (c) A solid conducting sphere of the same radius has charge $+9.14\times10^{-5}$ C. What is the strength of the electric field 2.50 m from the center of the sphere? (d) What is the direction of the field?

(a) __________ N/C
(b) i. Directly towards the center of the sphere
ii. Radially out from the sphere
(c) __________ N/C
(d) i. Directly towards the center of the sphere
ii. Radially out from the sphere

9.3 A conducting sphere of radius 23.0 cm has an unknown positive charge distributed uniformly over its surface. The strength of the electric field at a point 0.560 m from the center of the sphere is 12.0 N/C. What is the charge?

__________ C

9.4 An alpha particle (m = 6.68×10^{-27} kg, q = 3.20×10^{-19} C) circularly orbits a charged sphere at a radius of 5.00 cm from the sphere's center. The electric attraction between the two objects provides the centripetal force to keep the particle in orbit. The orbital speed of the particle is 6.56e+4 m/s. What is the charge on the sphere?

__________ C

9.5 There is a charged spherical shell, with +3.00 mC uniformly distributed over its surface. At its exact center is a point charge of −3.00 mC. (a) What is the strength of the electric field at any point outside the shell? (b) What is the strength of the electric field at any point inside the shell, a distance r from the center? (c) Draw a (two-dimensional) diagram of the field.

(a) ○ 0 N/C ○ $(6.00\times10^{-3})k/r^2$ ○ $(3.00\times10^{-3})k/r^2$
(b) ○ 0 N/C ○ $(6.00\times10^{-3})k/r^2$ ○ $(3.00\times10^{-3})k/r^2$
(c) Submit answer on paper.

9.6 A solid, nonconducting sphere of radius R has a positive charge Q distributed uniformly throughout its volume. Since the sphere is nonconducting, the charge does not migrate to its surface. (a) What is the volume charge density throughout the sphere? (b) Construct a spherical Gaussian surface of radius r inside the sphere, where $r \le R$. How much charge is contained inside the Gaussian surface? (c) What is the area of the Gaussian surface of radius r? (d) The electric field points radially outward at all points within the sphere (and outside too, for that matter). What is the strength of the electric field at a distance r from the center, inside the sphere?

(a) ○ $Q/\pi R^2$ ○ $3Q/4\pi R^3$ ○ $4Q/3\pi R^2$
(b) ○ Qr/R ○ Qr^2/R^2 ○ Qr^3/R^3
(c) ○ $4\pi r^2$ ○ $4\pi r^2/3$ ○ $4\pi r^3/3$
(d) ○ $Qr/4\pi R^3\varepsilon_0$ ○ $4\pi R^3 Q\varepsilon_0/r$ ○ $Qr^2/4\pi R^2\varepsilon_0$

Section 10 - Electric field of two infinite planes: a capacitor

10.1 A homemade capacitor is built by placing two flat steel cookie sheets next to each other, with thin rubber spacers in between, so that they are separated by 1.25 mm. One sheet is given a positive charge of 7.93 μC, and the other is given an equal and opposite negative charge. Each sheet measures 35.0 cm by 45.0 cm. (a) What is the strength of the electric field between the sheets? (b) If extra spacers are inserted to double the separation between the sheets, what is the new electric field strength? (c) If the sheets are moved a kilometer apart, the electric field between them becomes essentially zero. Explain the discrepancy between this result and the results observed in parts "a" and "b" of this problem.

(a) __________ N/C (b) __________ N/C

Additional Problems

A.1 The flux passing through a part of the surface of a sphere, of area 0.010 m^2, is 4.0 (N/C)·m^2. The sphere has radius 5.0 cm and is centered on a point charge. There are no other charges present. Determine the charge enclosed by the sphere.

__________ C

27.0 - Introduction

We began our study of electricity and magnetism with electrostatics, which focuses on the forces and fields created by stationary charges.

Now we will concentrate on charges that are in motion, a branch of physics called electrodynamics. We will discuss the flow of charge (electric current), resistance to current, and the relationship among potential difference, resistance, and current.

In this chapter, we focus primarily on the fundamentals of these topics rather than their applications. We will not concern ourselves too much about the source of the potential difference, nor worry too about where the current may be flowing or why one would want it to flow. Once we have discussed the essentials, we can apply these concepts to electric circuits, and consider typical sources of potential difference, such as batteries, and common sources of resistance such as resistors and light bulbs.

27.1 - Electric current

Electric current: Amount of net charge passing through a surface per second.

Current is the rate at which net charge passes through a hypothetical surface, the number of coulombs of charge per second. As a practical matter, currents are often measured as they pass through a wire, so we use this configuration to explain currents in the illustrations to the right. In this section, we focus solely on the current, not its cause.

To measure the amount of current, we can place an imaginary surface across the wire shown in Concept 1 and count the net number of electrons moving through it each second. The electrons we show are the *charge carriers*, the charges that make up the current. In this example, they are moving from right to left.

Counting the electrons as they pass by is a useful start, but electric current is charge per second, so we need to multiply by the charge of an electron. An electron has a charge of -1.6×10^{-19} C, so if there are five electrons flowing by every second, the electric current is 8.0×10^{-19} coulombs per second. For reasons discussed below, the flow of current here is considered positive.

The equation in Equation 1 states that current is the net charge passing through a surface divided by time. The *ampere* (A) is the unit for electric current. It is named after the French scientist André-Marie Ampère. One ampere equals one coulomb per second. A coulomb of charge is equivalent to 6.2×10^{18} electrons, so one ampere equals that number of electrons passing through every second. The letter I represents current.

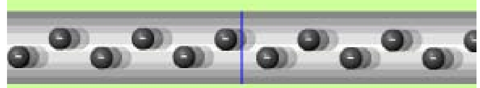
Using water as an analogy may help you understand electric currents. The rate of water flow is measured in several settings. You may have seen water flow ratings for shower heads. Newer shower heads allow a water flow of nine liters per minute, about half the rate of flow of older models. Boaters also keep a close eye on water flow. For a particular river, a rafter might measure how much water is flowing, in cubic feet per second, to judge whether it is safe to run the rapids.

Similarly, the rate of flow of electrons can be measured. Many electrons move in the currents used in household devices. About 2.5×10^{18} electrons pass through the light bulb in a typical household flashlight every second, which is 0.4 amperes.

Current is a scalar quantity. It states the rate of flow of charge, not the direction of the flow of charge. Again, think of water. Liters per second tells you how much water is flowing, but not in which direction. Often, it is important to know the direction of an electric current: which way the charge is flowing. Confusingly, the arrow used to indicate a current's direction points opposite the way you might expect. It does not point in the direction that electrons flow. Instead, it points the way positive charges would be flowing if they were the charges moving in an electric current. This is shown in Concept 2. The arrow indicates the direction of what is called *conventional current*, the direction in which positive charge carriers creating the current would move. The flow of electrons discussed above would be described as a positive current flowing to the right, not a negative current flowing to the left.

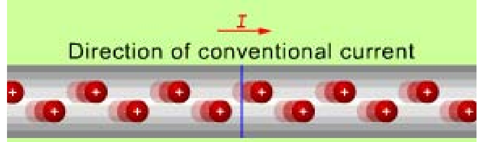
Why is typical electric current shown as though positive charge carriers flow, even if they do not? Scientists began studying electric currents before they knew about electrons and protons. One of the early explorers of electricity, Benjamin Franklin, established the convention that the current points in the direction of positive charge flow. More than a hundred years later, when the scientist Edwin Hall determined that

concept 1



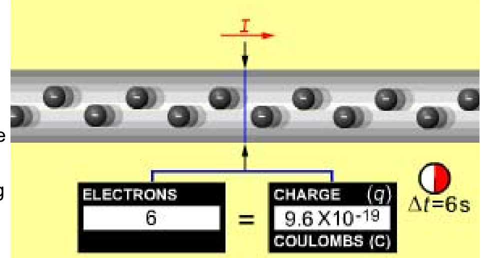
Electric current
Rate of net flow of charge
Typically composed of moving electrons

concept 2



Direction of conventional current
Current arrow in direction of positive charge flow

equation 1



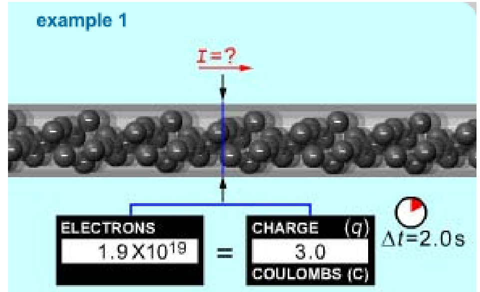
Electric current

$$I = \frac{\Delta q}{\Delta t}$$

current is actually the movement of negative charge carriers (electrons), the positive charge flow convention had already been established, and it remains in place to this day.

In basic configurations like the segment of conducting wire shown to the right, electric current is stated as a positive quantity and its direction is indicated with an arrow. In other circumstances, such as alternating current circuits where the direction of current flow changes constantly, signs can be used to indicate the direction of current.

I = current
Δq = charge passing through surface
Δt = elapsed time
Units: amperes (A)



In 2.0 seconds, 3.0 coulombs of charge flow past this charge counter. What is the current?

$$I = \frac{\Delta q}{\Delta t}$$

$$I = \frac{3.0\ \text{C}}{2.0\ \text{s}}$$

$I = 1.5$ A

27.2 - Drift speed

Drift speed: Average speed of charge carriers that make up a current.

Whether or not there is an electric current in a wire, the electrons in the wire move randomly at high speed, on the order of a million meters per second, due to their thermal energy. This random motion does not create a current because it causes no **net** motion of charge over time.

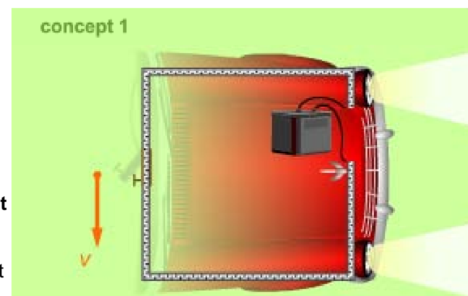
When there is an electric current, there is a net motion of charge. The average speed at which charge moves in a current is called its drift speed. To illustrate this concept, we use the simplified diagram of the wiring in a car shown in Concept 1. In the diagram, a battery is shown; the potential difference between its terminals causes electrons to move counterclockwise along the wire. The average speed at which they move along the wire is the drift speed.

The magnitude of the drift speed may surprise you; it is on the order of 10^{-4} m/s. That is less than a meter per hour, slower than a snail slimes across the ground. This may seem surprising because when you flip the car's headlight switch, the lights illuminate almost instantaneously. It may seem as if electrons are flowing from near the battery to the headlights at an incredible speed.

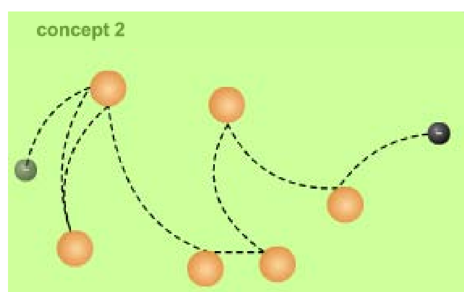
Why do the lights go on so quickly? When the switch is flipped, the electric field created by the car's battery travels through the wires at nearly the speed of light, 3.00×10^8 m/s. It causes electrons everywhere in the wire, including those in the filaments of the headlights, to start to move almost immediately. It is like when you turn on a water faucet: water flows out very quickly. You receive water near the faucet; you do not have to wait for water to flow all the way from the reservoir to your tap.

Why do the electrons move along the wire so slowly? If the electrons were in a vacuum, the electric field would accelerate them until they moved at near the speed of light, since there would be no force opposing their motion. But in a wire, the electrons collide with the atoms that make up the wire. The combination of the collisions and the electric field cause them to move in the erratic zigzag pattern you see in Concept 2. (We overstate the curved element of their paths to emphasize the effect of the field on their motion.)

The motion of these electrons is like that of a ball in play in a pinball machine. The ball moves fairly randomly as it rebounds from bumpers and other obstacles. However, the surface of a pinball table slants downward, so despite the erratic motion caused by the bumpers, a component of the gravitational force pulls the ball down toward you over time. (In short, the maker of the pinball game uses gravity to ensure that the ball "drifts" downward and "drains".)



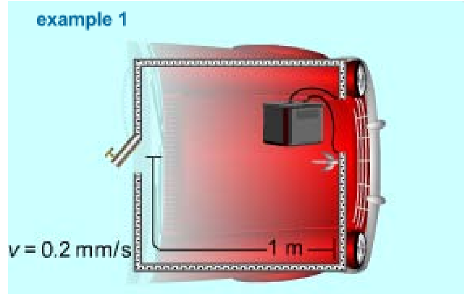
Drift speed
Average speed of charges in current
In typical wiring, about 0.1 mm/s



Electrons zigzag ...
... but move along wire

There is no simple relationship between drift speed and the amount of current in various conductors. A great number of charge carriers moving slowly can result in a larger electric current than a smaller number of charge carriers flowing quickly.

There is a final point to be made about the erratic motion of electrons in the current. Energy transfers from the electrons to the atoms during the collisions. The collisions increase the random (thermal) motion of the atoms. This explains why the temperature of the filaments in the car headlights increases when a current flows through them: Energy is being transferred from the electrons in the current to the atoms that make up the filaments. The filaments become hot and emit light. The other parts of the wiring increase in temperature as well, but the system is designed to maximize the temperature increase in the filaments and minimize it elsewhere.



If the switch is 1.0 meter away from the headlights, and charge carriers move at the speed shown above, how long will it take for the headlights to light once the switch is thrown?

Nearly instantaneous

· Field acts along entire wire very quickly

27.3 - Ohm’s law and resistance

Resistance: The ratio of the potential difference across a conductor to the current through it.

Resistor: An electrical component often used to control the amount of current flow.

Resistors wired into a computer circuit board. Each resistor has a color code that indicates its resistance in ohms.

Resistance is defined as the potential difference across two points on a conductor divided by the current flowing through the conductor. At the right, we use a common electrical component called a resistor to illustrate this concept. There is a potential difference across the resistor and current flowing through it. Divide the potential difference by the current and you have calculated the resistance of this resistor.

The resistance of a resistor such as those shown above or to the right is constant. The resistor is made of *ohmic* materials, which are empirically known to obey Ohm’s law. Increase the potential difference and the current increases proportionally. The resistance does not change. The linear relationship between potential difference and current is shown as the first equation in Equation 1.

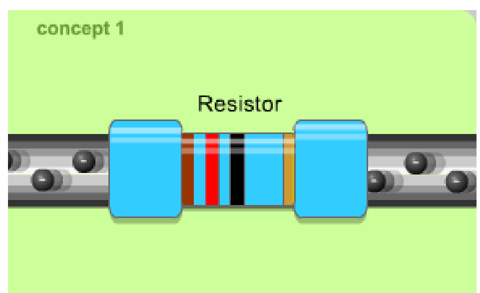
Since resistance equals potential difference divided by current, its unit, the ohm, is volts per ampere. Resistance is represented by Ω, the Greek letter omega.

Resistors are not the only components that resist the flow of current. The filament in a light bulb or the coils of an electric hot plate both function as resistors. The term “resistor” broadly refers to any element that is a source of resistance.

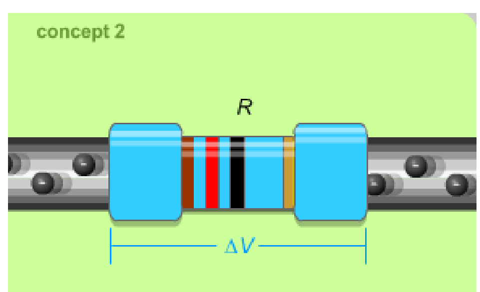
Again, water is a good analogy. The potential difference driving current in a wire resembles the pressure exerted on water in a pipe. Increase the pressure, and the water flow increases. Different pipes have different amounts of resistance. For instance, one with a rough interior wall would have greater resistance to water flow than one with a smooth wall.

Materials that do not obey Ohm’s law are called *non-ohmic*. Many components used in modern circuitry are made of non-ohmic materials. For instance, a diode has little resistance to current flow in one direction, and great resistance to current flow in the other.

Georg Ohm published his major work, including what we now know as Ohm’s law, in



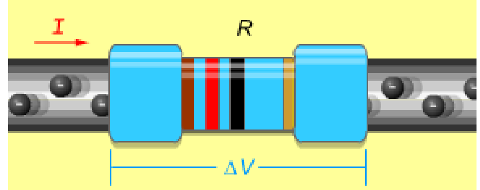
Resistor
Component used to regulate current



Resistance
Potential difference divided by current

1827. His theories were greeted with skepticism and his career was slow to progress. Why it took so long for his work to be appreciated is hard to say. Perhaps it is because the law is empirical as opposed to a fundamental law of nature. It is fair to note that many major leaps forward in physics were met with skepticism and opposition.

equation 1



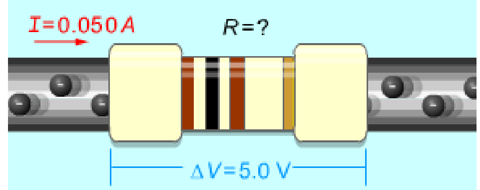
Ohm's law: potential difference and current

$$\Delta V = IR$$

$$R = \frac{\Delta V}{I}$$

ΔV = potential difference
I = current
R = resistance
Units of resistance: ohms (Ω), volts/ampere

example 1



What is the resistance of this resistor?

$$R = \frac{\Delta V}{I}$$

$$R = \frac{5.0\text{ V}}{0.050\text{ A}}$$

$R = 100\ \Omega$

27.4 - Interactive problem: Ohm's law

The interactive problem on the right shows a tanning bed, an electrical appliance through which current can flow.

Your challenge is to configure the tanning bed so that 2.4 amperes of current flow through it. If you accomplish this goal, Joe Pasty will get a perfect tan. Anything less and he will remain pasty. Anything more and it will be red and toasty for Joe.

In the simulation, the potential difference applied to the lamps of the tanning bed is 240 volts. (Assume that this potential difference is unchanging.)

You can control the tanning bed's resistance using the up and down buttons in the control panel. The resistance can vary from 10 to 200 ohms.

Use Ohm's law to calculate the tanning bed resistance that will result in 2.4 amperes of current flowing through the device. Set the resistance to the nearest 10 ohms, then press GO to see if you were correct. A gauge located on the tanning bed will display the amount of current flowing through the device.

interactive 1

Send a 2.4 A current through the tanning bed to give Joe Pasty the perfect tan.

27.5 - Resistivity

Resistivity is used to quantify how much a material resists the flow of current. It is represented by the Greek letter ρ (rho). A material considered a good conductor, such as copper, has low resistivity. Materials used as insulators, such as glass or rubber, have high resistivity.

Resistivity is the inverse of conductivity (which quantifies how well a material conducts current). Current does not readily flow through materials with very high resistivity. As the table in Concept 2 shows, there is an enormous range in the resistivity of materials.

The table includes nichrome, an alloy of nickel and chromium. Due to its high resistivity, Nichrome is commonly used in hot plates, toasters and other electrical appliances that generate intense heat. Its large amount of resistivity causes electrons in the current to lose energy to the nichrome atoms, increasing their thermal energy.

Extension cords, designed to allow current to flow safely from one point to another, take advantage of the relative resistivities of materials. The cords typically are composed of copper wire, a material with low resistivity, surrounded by rubber or vinyl, materials with high resistivity. Electrical current flows readily through the copper, but not through the material that surrounds it.

Resistivity is a property of a material. Its units are the ohm·meter (Ω·m). In contrast, resistance is the property of a component like a resistor or the filament in a light bulb. Resistivity is one factor that determines a resistor's resistance; the other factor is the geometry of the resistor. The longer a resistor, the greater its resistance; the wider it is (the greater its cross sectional area), the less its resistance. Again, water can be used to provide an analogy: a long, thin pipe resists the flow of water more than a short, fat one.

The equation used to calculate resistance as a function of a resistor's geometry and composition is shown in Equation 1.

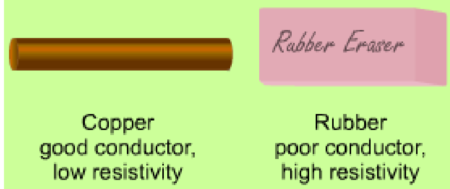
The problem in Example 1 asks you to compute the amount of resistance found in a segment of copper wire. The resistance of resistors that are sold in electronics shops runs from the tens of ohms to the thousands of ohms and beyond. When electric circuits are analyzed, the resistance of wire is often ignored because it is relatively insignificant.

The low resistivity of wire, along with Ohm's law, helps to answer a question physicists like to pose: Why can a bird safely stand on an unshielded high voltage line? Why does the bird not get injured as current flows up one of its feet, through its body, and down the other foot?

The answer is that the potential difference is very small across the segment of wire involved because the wire's resistance is quite low. Its resistance is low because the segment is wide and short and is made of a material with low resistivity. Ohm's law states that the potential difference is the product of the current and the resistance. With low resistance, there is very little potential difference. With very little potential difference, effectively no current flows through the bird.

Do NOT test this out yourself. High voltage lines are enormously dangerous. If you touch such a line while standing on the ground, you are providing a path for current to flow from the wire to the Earth. The potential difference between these points is enormous, and the resulting current could kill you.

concept 1



Resistivity
Inverse of conductivity
· Conductors have low resistivity
· Insulators have high resistivity

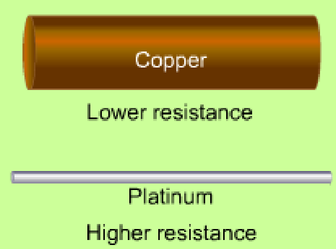
concept 2

RESISTIVITIES (Ω · m), at 20° C

Conductors		Semi-conductors	
Silver	1.6×10^{-8}	Carbon	3.5×10^{-5}
Copper	1.7×10^{-8}	Silicon	2.5×10^{3}
Aluminum	2.7×10^{-8}		
Iron	9.6×10^{-8}	**Insulators**	
Platinum	10.5×10^{-8}	Glass	10^{10} - 10^{14}
Nichrome	107.5×10^{-8}	Rubber	1.0×10^{13}

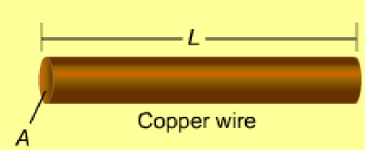
Table of resistivities, ρ

concept 3



Resistance
Depends on resistivity, configuration of material

equation 1

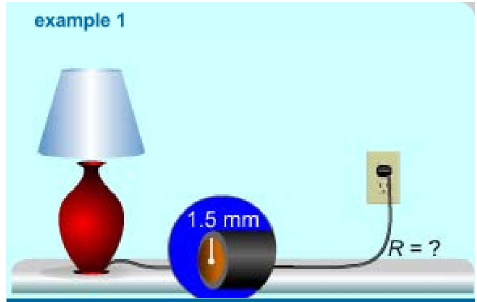


Resistance

$$R = \rho \frac{L}{A}$$

R = resistance of component
ρ = resistivity of material
L = length of component
A = cross-sectional area
Units of resistivity: Ω·m (ohm-meters)

example 1



The lamp cord is 0.75 meters long and contains copper wire. What is the wire's resistance?

$$R = \rho\frac{L}{A} = \rho\frac{L}{\pi r^2}$$

$$R = (1.7\times 10^{-8}\ \Omega\cdot\text{m})\ \frac{(0.75\ \text{m})}{\pi(0.0015\ \text{m})^2}$$

$R = 0.0018\ \Omega$

27.6 - Interactive checkpoint: extension cord

Extension cord wires come in different diameters, referred to as their gauge. For historical reasons, higher gauge numbers correspond to thinner wires. You have an extension cord with 12-gauge wires that is 15.0 meters long. A 12-gauge wire has a diameter of 2.05 mm. What is the resistance of a wire in the extension cord? (Use copper's resistivity at room temperature, which is 1.68×10^{-8} Ω·m.)

Answer:

R = ______ Ω

27.7 - Resistivity and temperature

The resistance of many conductors increases with temperature. As the temperature of the metal coils of the hot plate shown on the right increases, so do the resistivity of the metal and resistance of the coils. Other materials, such as semiconductors, have decreasing resistivity with temperature.

Equation 1 shows two equations that reflect the relationship of resistivity and resistance to temperature. Both include the *temperature coefficient of resistivity*, represented by α (the Greek letter alpha). A table of these coefficients is shown in Concept 2 for some materials. These coefficients are empirically determined and apply over a specific range of temperatures. To apply the resistivity equation, the material's resistivity must be known at one temperature, T_1. Its temperature coefficient of resistivity must also be known for the temperature T_1. These values are used to calculate the resistivity at another temperature.

The second equation is used to determine the change in resistance of a resistor. If the

concept 1

In many materials:

resistance of a resistor is known at one temperature, the equation can be used to calculate its resistance at another temperature. This equation can be derived from the first.

In Example 1, we use this equation to calculate the change in resistance of a nichrome coil in a hot plate that is heated from 25°C to 375°C.

Resistivity and resistance vary linearly with temperature

concept 2

	Temperature coefficient of resistivity, α, at 20°C (°C^{-1})
Brass	0.0015
Copper (drawn)	0.00393
Iron (electrolytic)	0.0064
Platinum silver	0.00031
Silver	0.0038
Tungsten	0.005
Carbon	negative*
Germanium	negative*
Silicon	negative*

*Dependent upon purity of sample

Temperature coefficient of resistivity, α

Positive α indicates resistivity increases with temperature

Negative α indicates resistivity decreases with temperature

equation 1

Temperature coefficient of resistivity, α

$$\rho_2 = \rho_1[1 + \alpha(T_2 - T_1)]$$
$$R_2 = R_1[1 + \alpha(T_2 - T_1)]$$

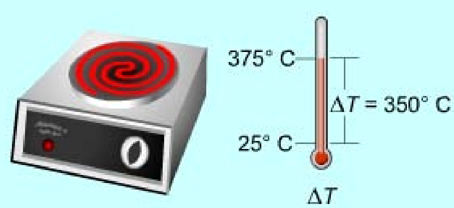
ρ = resistivity

α = temperature coeff. of resistivity

T = temperature

R = resistance

example 1



The hotplate contains a nichrome wire with a resistance of 15 Ω at 25°C. What is its resistance at 375°C? For this nichrome alloy, α = 4.0×10^{-4} °C^{-1} at 25°C.

$R_2 = R_1[1 + \alpha(T_2 - T_1)]$

$R_2 = R_1[1 + (4.0\times10^{-4}\ °C^{-1})(T_2 - T_1)]$
$R_2 = 15\ \Omega\ [1 + (4.0\times10^{-4})(350)]$
$R_2 = 17\ \Omega$

27.8 - Electric power

Power is defined as work per unit time. Power also can be stated as the amount of energy transfer per unit time, and this is often a useful formulation when considering electric circuits.

Electrical devices are often rated on how much energy they use per second. A 100-watt light bulb requires an average of 100 joules every second. As it operates, the light bulb's filament warms up and emits heat. This flow of heat from a resistor is called *joule heating*.

Power can also be calculated as the product of current flowing through a resistor and the potential difference across it. This is the equation in the first line in Equation 1. Ohm's law can be used to then restate this equation in terms of resistance and current, or in terms of potential difference and resistance. We show these equations in the second line in Equation 1.

In the example problem, we simplify things by using an average or constant value for current and potential difference − likely this burner is powered with an alternating current where these values vary. Alternating current (AC) is a topic later in the textbook.

Below we derive the first equation, $P = I\Delta V$. Before deriving it, we review why a resistor consumes power. The essential component of many household devices − toasters, light bulbs, electric burners − is a resistor.

As electrons move through any resistor, they collide with the atoms that make up the resistor. The electrons lose energy in these collisions and the atoms gain it, which increases the temperature of the resistor. More current through a given resistor means more collisions, and a warmer resistor. This analysis confirms one aspect of the power equation: Power increases with the amount of current.

To derive the power equation, we will consider the work done on the electron by the electric field. We will state the work in terms of the potential difference: It equals the charge times the potential difference. We will use that equation to derive the power equation.

Variables

work	W
total charge moving across resistor	q
potential difference across resistor	ΔV
time interval	Δt
power consumed by resistor	P
current passing through resistor	I

Strategy

1. State the equation that relates work, charge and potential difference.
2. Divide by the time interval Δt to calculate the work per unit time, which is power.

Equations

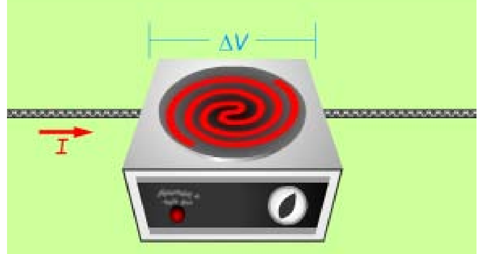
Here, work is related to charge and potential difference by

$$W = \Delta PE = q\Delta V$$

We will also use the definitions of power and current.

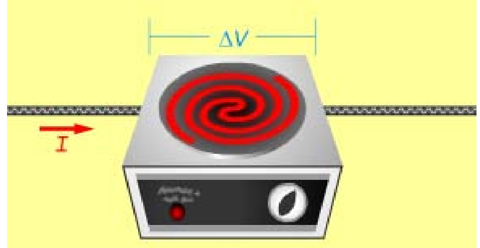
$$P = \frac{W}{\Delta t},\quad I = \frac{q}{\Delta t}$$

concept 1



Electric power
Function of current, potential difference

equation 1



Electric power

$P = I\Delta V$

$P = I^2R = \Delta V^2/R$

P = power
I = current
ΔV = potential difference
R = resistance

example 1

How much power is being supplied to this hot plate?
$P = I\Delta V$
$P = (8.00\ A)\ (120\ V)$
$P = 960\ W$

Step-by-step derivation

We state the relationship between work and potential difference, then divide by the time interval Δt. This converts work to power and total charge to current, resulting in the first formulation of the power equation.

Step	Reason
1. $W = q\Delta V$	work and potential difference
2. $\frac{W}{\Delta t} = \frac{q\Delta V}{\Delta t}$	divide by Δt
3. $P = I\Delta V$	definitions of power and current

Ohm's law is used to derive the other equations for power in the second line of Equation 1. For instance, $\Delta V = IR$, so we can substitute for ΔV and conclude that $P = I^2R$. The equation $P = I\Delta V$ holds true for all electrical devices, while the other equations apply solely to power (energy) dissipated as heat by resistors.

27.9 - Sample problem: solar panels

This public school uses solar power. Using information supplied below, determine how many watt-hours of electricity the solar panels will provide during their lifetime.

What will the cost per kW·h of the solar panel system be over its lifetime? Suppose the cost of traditional energy will average \$0.11 per kilowatt-hour over the next 25 years. Will the school save money?

A school in Los Angeles is deciding whether or not to install solar panels on its roof. The photovoltaic system under consideration provides 680 watts of power under direct sunlight, costs \$3500, and is expected to last 25 years, according to the experts at Stewart Solar Systems. In Los Angeles the average *insolation*, or equivalent hours of direct sunlight reaching the Earth's surface, is 5.62 hours per day.

(a) How many watt-hours of electricity will the solar panels provide during their lifetime? (A watt-hour is a unit of energy just like the Joule; it is calculated by multiplying the power and the amount of time that power is produced.)

(b) Suppose the cost of electricity coming from the local utility company is predicted to average \$0.11 per kilowatt-hour (kW·h) over the next 25 years. What is the cost per kW·h of the solar panel system over its lifetime? Should the school buy the solar panels?

Variables

generation power of solar panel system	$P = 680$ W
lifetime of system	$t = 25$ y
average equivalent hours of sunlight per day	$s = 5.62$ h/d
equivalent hours of sunlight over system lifetime	S
energy produced over system lifetime	E
price of utility electricity per kW·h	$p = \$0.11$
one-time cost of solar panel system	$C = \$3500$
cost of solar panel system per kW·h	c

What is the strategy?

1. Find the total number of hours of direct sunlight the solar panels will receive over their lifetime.
2. Calculate the number of watt-hours the solar panels will produce over their lifetime. Convert to kW·h.
3. Calculate the price per kW·h from the solar panels and compare it with the utility company's price.

Physics principles and equations

Power stated as energy used over a period of time

$P = \Delta E/\Delta t$

Step-by-step solution

First, find the total amount of energy the solar panels will produce over their lifetime.

Step	Reason
1. $S = st$	hours of sunlight over lifetime
2. $S = (5.62\ \text{h/d})(25\ \text{y})(365\ \text{d/y})$ $S = 51{,}283\ \text{h}$	enter values
3. $E = PS$	energy produced over lifetime
4. $E = (680\ \text{W})(51{,}283\ \text{h}) = 3.49 \times 10^7\ \text{W·h}$	enter values

For part (b), calculate the cost per kW·h.

Step	Reason
5. $E = (3.49 \times 10^7\ \text{W·h})/(1000\ \text{kW/W})$ $E = 3.49 \times 10^4\ \text{kW·h}$	convert to kW·h
6. $c = C/E$	cost per kW·h
7. $c = (\$3500)/(3.49 \times 10^4\ \text{kW·h})$ $c = \$0.10/\text{kW·h}$	enter values

This is **less** than the predicted cost of electricity from the utility company, so the school will save money in electricity costs if they install the solar panels. Even better, by using renewable energy they will contribute to a less polluted future.

On the web, you can find out more about solar energy and find out the average insolation anywhere in the world.

27.10 - Interactive checkpoint: running a dishwasher

Suppose you measure the current flow to a dishwasher and find that the average current is 9.50 A over the 45.0-minute dishwasher cycle. The dishwasher is powered by a potential difference of 120 V. If electricity costs \$0.0900 per kilowatt-hour, what is the cost of running one dishwasher load?

Answer:

$p = \$$ ______ per dishwasher cycle

27.11 - Sample problem: using a hair dryer

The label on an electric blow dryer says it uses 1100 W and is designed for 120 V outlets. How much current flows through the hair dryer? What is its resistance?

Here, you will treat the hair dryer as though it were a direct current (DC) appliance.

Variables

power rating	$P = 1100$ W
potential difference	$\Delta V = 120$ V
current	I
resistance	R

What is the strategy?

1. Use the equation for power as a function of current and voltage to find the current through the hair dryer.
2. Use Ohm's law to find the hair dryer's resistance.

Physics principles and equations

Power in an electric circuit

$$P = I\Delta V$$

Ohm's law

$$\Delta V = IR$$

Step-by-step solution

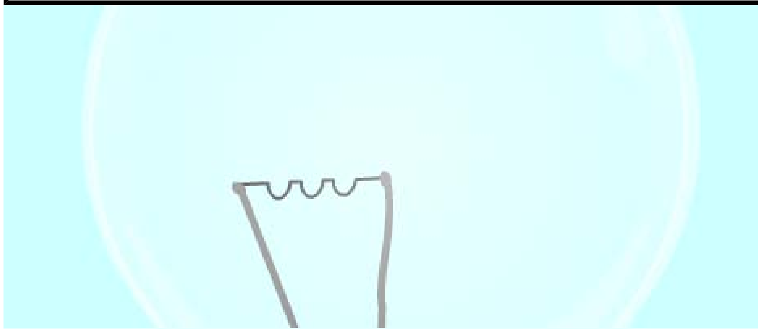
We use the power equation to find the current through the hair dryer.

Step	Reason
1. $P = I\Delta V$	power equation
2. $I = \dfrac{P}{\Delta V}$	solve for I
3. $I = (1100\text{ W})/(120\text{ V}) = 9.2\text{ A}$	evaluate

Now we can use the current to find the resistance of the hair dryer.

Step	Reason
4. $\Delta V = IR$	Ohm's law
5. $R = \dfrac{\Delta V}{I}$	solve for R
6. $R = \dfrac{120\text{ V}}{9.2\text{ A}} = 13\ \Omega$	evaluate

27.12 - Sample problem: light bulb filament



Here you see a 60-watt, 120-volt light bulb. The coiled tungsten filament of the bulb is 2.0 m long, and it heats to 3000° C during normal operation.

What is the cross-sectional area of the filament?

The extremely thin filament is 2.0 meters long and coiled so that it fits inside the bulb. The resistivity of tungsten is $5.28 \times 10^{-8}\ \Omega\cdot\text{m}$ at room temperature $(20°\ \text{C})$, and its temperature coefficient of resistivity is $0.0045\ \text{C}°^{-1}$. Assume the bulb is in a circuit that supplies a 120 V potential difference across the light bulb. Here, you will treat the light bulb as though it were a direct current (DC) device.

Variables

resistivity of tungsten, room temperature	$\rho_1 = 5.28\times10^{-8}\ \Omega\cdot\text{m}$
resistivity of tungsten, operating temp	ρ_2
room temperature	$T_1 = 20°\text{C}$
operating temperature	$T_2 = 3000°\text{C}$
temperature coefficient of resistivity	$\alpha = 0.0045\ \text{C}°^{-1}$
length of filament	$L = 2.0\ \text{m}$
cross sectional area of filament	A
resistance of filament, operating temp	R
potential difference across filament	$\Delta V = 120\ \text{V}$
power used by light bulb	$P = 60\ \text{W}$

What is the strategy?

1. Use the power equation to calculate the resistance at $3000°\text{C}$.
2. Use the temperature coefficient of resistivity equation to find the resistivity at $3000°\text{C}$.
3. Finally, use the resistivity equation to calculate the filament's cross sectional area.

Physics principles and equations

Resistance as a function of resistivity and geometry

$$R = \rho\frac{L}{A}$$

Resistivity as a function of temperature

$$\rho_2 = \rho_1[1 + \alpha(T_2 - T_1)]$$

Power in an electric circuit

$$P = I\Delta V = (\Delta V)^2/R$$

Step-by-step solution

First we calculate the operating resistance.

Step	Reason
1. $P = (\Delta V)^2/R$	power equation
2. $R = (\Delta V)^2/P$	solve for R
3. $R = (120\ \text{V})^2/(60\ \text{W}) = 240\ \Omega$	evaluate

Next we find the resistivity at operating temperature. The temperature difference is 2980 C°, which is 3000°C − 20°C.

Step	Reason
4. $\rho_2 = \rho_1[1 + \alpha(T_2 - T_1)]$	equation stated above
5. $\rho_2 = (5.28\times 10^{-8}\ \Omega\cdot\text{m}) \cdot [1 + (0.0045\ \text{C}^{\circ -1})(2980\ \text{C}^\circ)]$	enter values
6. $\rho_2 = 76.1\times 10^{-8}\ \Omega\cdot\text{m}$	evaluate

We use the resistance and resistivity values at 3000°C to find the cross-sectional area.

Step	Reason
7. $R = \rho_2 L/A$	resistivity and geometry equation
8. $A = \rho_2 L/R$	solve for A
9. $A = \dfrac{(76.1\times 10^{-8}\ \Omega\cdot\text{m})(2.0\ \text{m})}{240\ \Omega}$	enter values
10. $A = 6.3\times 10^{-9}\ \text{m}^2$	evaluate

27.13 - Sample problem: power transmission

A power plant provides 450 MW of power through transmission lines at a potential difference of 500 kV. What current flows through the lines?

For a city 300 km away from the power plant the total resistance in the transmission lines is 60 ohms. What fraction of the generated power is dissipated as “line loss”?

If the generated power is transmitted at a potential difference of 175 kV instead, what fraction is dissipated?

Variables

generated power	$P = 450\times 10^6$ W
transmission line potential difference	ΔV
resistance	$R = 60\ \Omega$
dissipated power	P_D

What is the strategy?

1. Find the current in the wire using the equation for power in terms of the potential difference and the current.
2. Calculate the power dissipated by the power lines using the equation for power dissipated by a resistor.
3. Divide the dissipated power by the generated power to find the fraction of the power dissipated by the lines.
4. Repeat this calculation for the last question asked above.

Physics principles and equations

Power as a function of potential difference and current

$$P = I\Delta V$$

Power dissipated by a resistor

$$P = I^2R$$

Power in an electric circuit

$P = (\Delta V)^2/R$

Step-by-step solution

First we calculate the current in the transmission lines when the potential difference is 500 kV. We do this using the power for the power plant and the potential difference across the wires, which are both stated above.

Step	Reason
1. $P = I\Delta V$	power equation
2. $I = \frac{P}{\Delta V}$	solve for I
3. $I = \frac{(450\times 10^6\text{ W})}{(500\times 10^3\text{ V})}$ $I = 900\text{ A}$	evaluate

First, we find the power dissipated as heat by the wire. Then we divide that by the amount of power generated by the plant to determine what fraction is wasted.

Step	Reason
4. $P_D = I^2R$	power dissipated by a resistor
5. $P_D = (900\text{ A})^2(60\ \Omega)$ $P_D = 48.6\times 10^6\text{ W}$	evaluate
6. $\frac{P_D}{P} = \frac{48.6\times 10^6\text{ W}}{450\times 10^6\text{ W}}$ $\frac{P_D}{P} = 0.108$	divide by generated power

For the second case we repeat the same calculations with a lesser potential difference.

Step	Reason
7. $I = \frac{450\times 10^6\text{ W}}{175\times 10^3\text{ V}}$ $I = 2570\text{ A}$	evaluate equation in step 2
8. $P_D = (2570\text{ A})^2(60\ \Omega)$ $P_D = 397\times 10^6\text{ W}$	power dissipated by resistor
9. $\frac{P_D}{P} = \frac{397\times 10^3\text{ W}}{450\times 10^3\text{ W}}$ $\frac{P_D}{P} = 0.882$	evaluate

The differing rates of power consumption explain why power companies transmit power at high potential differences. The potential difference is lowered at transformers near the consumer. (Transformers are discussed further in the chapter on electromagnetic induction.) The potential difference is lowered to 120 V, the standard potential difference of a power outlet in a house, just before it enters the house or building using the power.

27.14 - Gotchas

Do conventional current arrows reflect the flow of negative or positive charge? Positive charge. Although in a typical current found in household appliances and such, it is negatively charged electrons that flow, the convention is that the arrow indicates the direction of positive charge flow. If electrons flow to the left, the conventional current arrow points to the right.

Current decreases when passing through a resistor and then increases again upon exiting. No. The current before, in and after the resistor is the same.

A wire with a current has a net electrostatic charge. No. Although there is a net charge flowing by any given point, the overall wire is neutral.

A person walks by. The person contains electrons. Therefore, there is a current. No, the person is electrically neutral. Current is the flow of net charge. In this case, there is no movement of **net** charge and therefore no current. If the person were electrically charged, there would be a current as the person passed by.

27.15 - Summary

Electric current is the rate at which charge flows through a conductor. The symbol for current is I and it is measured in amperes (A), where 1 A = 1 C/s.

Current usually consists of moving electrons, which have negative charge. However, current is almost always represented in descriptions and diagrams as conventional current, which is a flow of positive charges that would constitute the same current, so the direction of conventional current is opposite to the actual movement of electrons. Even though drawings often show the direction of current with an arrow, current is a scalar.

Individual electrons in a wire travel much more slowly than the electric field that propels them. Collisions with the atoms that make up the wire prevent the electrons' continued acceleration, and cause them to follow a meandering zigzag path. The drift speed is the average net speed of electrons along the wire. In household wiring, it is about the speed of a snail.

Current density is the amount of current per unit area flowing through a conductor. Unlike current, it **is** a vector. Its magnitude is measured in A/m^2. One way to calculate current density is by dividing the current in a wire by the wire's cross sectional area. Another way is to multiply the number density of the charge carriers in the wire by their individual charge and by their drift speed. The direction of the current density vector equals the direction of the average drift velocity of the charge carriers, which are assumed to be positive. In other words, the direction of the current density vector is the same as the direction of the electric field.

For many materials (called "ohmic" materials), the current density is proportional to the applied electric field. Current density and electric field are related by the material's conductivity, a measure of how well the material conducts electrons.

Ohm's law states the relationship between potential difference, current and resistance. A resistor is an electrical component that can be used to regulate current. A particular resistor is characterized by its resistance. Resistance is measured in ohms (Ω). 1 Ω = 1 V/A.

Resistivity is a measure of how much the material resists current. For a simple wire made of a single material, the resistance is the resistivity of the material times the wire's length, divided by its cross-sectional area.

The temperature of a material can affect its resistivity.

Energy is dissipated, often as heat, when an electric current passes through a resistance. Electric power measures the amount of energy consumed per unit time. It is equal to the current through a circuit element times the potential difference across it. Sometimes, as when a battery is being charged, the energy consumed is stored rather than dissipated. For resistors that obey Ohm's Law, the two equations shown in the last line on the right may also be used to calculate power.

Equations

Definition of current

$$I = \frac{\Delta q}{\Delta t}$$

Ohm's law

$$\Delta V = IR$$

Resistance and resistivity

$$R = \rho \frac{L}{A}$$

Electric power equations

$$P = I\Delta V$$

$$P = I^2 R = \Delta V^2 / R$$

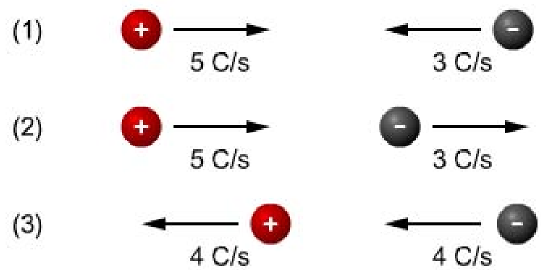
Chapter Assumptions

For any problems in this chapter that involve a household electrical outlet, assume that the voltage supplied is fixed at 120 V. While this is not true for actual outlets, whose voltage varies and periodically exceeds 120 V, the simplification will not affect the accuracy of your answer.

Conceptual Problems

C.1 The three parts of the illustration show positive and negative charges moving in the indicated directions, at the rates shown. (a) Which scenario depicts the greatest current? (b) Which scenario depicts the smallest current?

(a) ○ 1 ○ 2 ○ 3
(b) ○ 1 ○ 2 ○ 3



C.2 3 amperes of current flow through a wire. (a) What net charge flows past a point in the wire each second? (b) What is the net charge on the wire?

(a) ________ C
(b) ________ C

C.3 A current of 2 amperes flows to the left in a stationary wire. If the wire also starts moving to the left, what happens to the current as measured by a stationary observer? Why?

i. The current increases
ii. The current decreases
iii. The current stays the same

C.4 Two like charges of one coulomb each, separated by one meter, repel each other with a force of 8.99×10^9 N, equivalent to the weight of 10 battleships. A flashlight bulb may have a current of 0.40 A flowing through it, which is equivalent to one coulomb of charge flowing through every 2.5 seconds. But a flashlight circuit is clearly not subject to tremendous electrostatic forces -- or even small ones. Why not?

C.5 A section in the textbook that explains drift speed displays a diagram that shows the zigzag path followed by a black electron as it pursues an average path to the right through a section of conducting wire, but collides with many yellow atoms in the wire on its way. Suppose that the black circle in the diagram were a steel ball instead of an electron, and that the yellow circles were bumpers in a pinball game instead of atoms. That is, suppose this were a diagram of a pinball machine in play. In a pinball game, what field is acting analogously to the uniform electric field that propels electrons through a wire?

C.6 The section in the textbook that explains drift speed displays a diagram that shows the zigzag path followed by an electron as it pursues an average path to the right through a section of conducting wire, but collides with many atoms in the wire on its way. The diagram correctly shows the zigs and zags as being curved rather than straight line segments. (a) Why is this? (b) What is the shape of the curved line segments?

(a) (b)

C.7 A resistor has 2.0 A of current flowing through it when the potential difference across it is 6.5 V. When the potential difference is increased to 19.5 V, the current increases to 5.5 A. What kind of material is this?

○ Ohmic ○ Non-ohmic

C.8 Two wires are made from materials, such as copper and aluminum, that have different resistivities. The wires have the same length. Is it possible for them to have the same resistance? Explain.

○ Yes ○ No

C.9 Would you expect there to be any materials that have both high resistivity and high conductivity? Explain.

○ Yes ○ No

C.10 Equations are given in the textbook that allow you to calculate how the resistivity of materials and resistance of objects vary with the temperature. The discussion of the equations does not specify whether the temperatures in the equations should be measured in kelvin or in degrees Celsius. Does it matter? If it matters, deduce which should be used. If it doesn't matter, tell why not.

i. It matters
ii. It does not matter

C.11 You buy a light bulb in the U.S. labeled 75 watts. Since you bought it in the U.S., the power rating assumes that the potential difference across the light bulb will be 120 volts. Instead, you provide it with a potential difference of 220 volts (the potential difference provided by wall outlets in most European countries). Will the light bulb dissipate 75 watts of power? Explain.

○ Yes ○ No

C.12 Solar panels are priced based on the number of kilowatts they generate. You pay the utility an amount based on the number of kilowatt-hours you use. What is the connection between kW and kW·h?

C.13 You receive your bill from the electric company, and it states that you have used 3470 kW·h (kilowatt·hours) of electricity this month, and that due to the energy crisis in your town your bill is an astronomical $832.80. What is it, exactly, that you are paying for?

i. Electric power
ii. Electric energy
iii. Electric potential
iv. Some hotshot energy trader's vacation in Tahiti

Section Problems

Section 1 - Electric current

1.1 A current of 1.00 mA passes through a copper wire for 60.0 s. (a) How many coulombs of charge pass any point in the wire during this time period? (b) How many electrons pass any point in the wire during this time period? (c) How many protons pass any point in the wire during this time period?

(a) __________ C
(b) __________ electrons
(c) __________ protons

1.2 In a Crooke's Tube, a stream of electrons boils off a negative electrode, or cathode, and flies through a vacuum toward a positive electrode, or anode. Often, the anode is given a distinctive shape, like a cross, so that the electrons that miss it and fly past will create a dramatic shadow on the far end of the glass tube. The stream of electrons is flowing past a certain point between the cathode and the anode at a rate of 3.50 A. If 95.0% of the electrons reach the anode, (a) what amount of charge reaches the anode during a period of 15.0 s? Be careful of the sign of your answer. (b) How many electrons reach the anode during a period of 15.0 s?

(a) __________ C
(b) __________ electrons

1.3 A mouse sits next to a cell phone charger cord for 4.10 seconds, debating whether it is something worth chewing on. If the cord carries a current of 3.00 amperes, what is the magnitude of the charge that goes by the mouse while it thinks?

__________ C

1.4 According to the Bohr model, a hydrogen atom in its lowest energy state has a nucleus consisting of a single proton, which is orbited by a single electron. The speed of the electron is 2.19×10^6 m/s and the radius of its orbit (the "Bohr radiuis") is 5.29×10^{-11} m. What is the magnitude of the "current" flowing around the nucleus in the Bohr model?

__________ A

1.5 A wire spans two telephone poles. A tightrope walker picks up a sphere with a charge of 5.3 microcoulombs at one end and carries it to the other end in 7.2 seconds. What is the average current from one end of the wire to the other?

__________ A

1.6 A wire etched onto a computer chip carries a current of 4.1e−8 A. In 1.8 milliseconds, how many electrons flow past a point on the wire?

__________ electrons

1.7 A vacuum cleaner runs on a current of 12 amperes. How long will it take for 5.0 microcoulombs of charge to flow through it?

__________ s

1.8 Physicists often demonstrate the power of electricity by simulating lightning with a device called a Tesla coil. Charge is deposited onto a metal sphere until it amasses so much electric potential that it discharges through the air in the form of a spark. With larger Tesla coils, this spark resembles a miniature lightning bolt. Find the average current into a Telsa coil that amasses 1.38×10^{-6} C of charge in 5.72 s.

__________ A

1.9 The conducting sphere on top of a Van de Graaff generator is charged by a silk conveyer belt that picks up electrons at the bottom of the generator's column, from grounded brushes consisting of certain substances (for example, fur), and releases them at the top of the column to another set of brushes consisting of say, hard rubber. Suppose the belt of a certain Van de Graaff generator is 10.0 cm wide, and it travels upward at 45.0 m/s, carrying a "current" of 175 mA. (a) What is the surface charge density of the electrons on the belt? (b) Is the conventional "current" in the Van de Graaff generator flowing up or down?

(a) __________ C/m^2

(b) ○ Up ○ Down

1.10 Two wires are connected to a conducting sphere of radius 7.75 cm, which is initially uncharged. One wire carries a current of 3.47 μA into the sphere, and another wire carries a current of 1.26 μA out of the sphere. How long does it take to to produce an electric potential of 5.00 V at a distance of 11.6 cm away from the center of the sphere?

__________ s

1.11 Conducting objects can be chrome-plated by immersing them in an electrolytic bath having positive chromium ions in solution. Each ion carries a charge of $+e$, and the object to be plated, say an automobile bumper, is wired as the negative electrode in the bath. A positive electrode is also placed in the bath, and as a current flows between the two electrodes, chromium ions get deposited on the negative electrode (the bumper). If 0.186 kg of chromium is deposited on a bumper in 105 min, what is the current passing through the bath? The molar mass of chromium is 0.0520 kg.

__________ A

Section 2 - Drift speed

2.1 In one of Bill Cosby's comedy routines, he talks about being so tired that "I turned off the light switch, and I was in my bed before the light went out!" Suppose the drift speed of electrons in the wiring of the Cosby household is 3.0×10^{-4} m/s, and that there are 3.2 m of wiring between the light switch in his bedroom and the lamp. (a) In a direct current circuit, where the current flows one way, how long would it take a conduction electron to go from the switch to the lamp? (b) Even though Cosby couldn't really have gotten into bed before the light went out, could he at least have gotten into bed before one electron had traveled from the switch to the lamp?

(a) __________ s

(b) Submit answer on paper

2.2 In a typical household current, the electrons in the wire may have a drift speed of 1.3×10^{-4} m/s. Actual household current is not a direct current, but instead is an alternating (oscillating) current. (If you have ever received an electric shock from an outlet, you have felt this rapid alternation as a painful vibration. A safe way to see the alternation is to wave your fingers rapidly in front of a fluorescent light, and observe the stroboscopic stutter-silhouette that they form.) You can model the motion of each conduction electron as simple harmonic motion, with a frequency of 60 Hz. The drift speed is the maximum speed, what the speed would build up to if the electric field were applied continually, as in the case of direct current. What is the approximate amplitude of an electron's oscillation?

__________ m

Section 3 - Ohm’s law and resistance

3.1 A resistor has a potential difference of 5.65 volts applied across it, and a current of 345 mA flows through it. What is its resistance?

__________ Ω

3.2 A device found at some amusement parks asks you to "test your personal force" by grasping two conducting bars, one in each hand, and tolerating a muscle-clenching flow of direct current through your body. Assuming that your body is an ohmic material, the resistance of your body between your two hands (if you are not sweating!) is about $1.00\times10^4\ \Omega$. What voltage does the machine have to apply to the conducting bars to drive a current of 30.0 mA through your body, causing strenuous muscle contraction?

__________ volts

3.3 The resistance across your (wet) tongue, from one side to the other near the tip, is 375 Ω (assume that your tongue is an ohmic material). If you connect two thick copper wires of relatively negligible resistance to the terminals of a 1.5 volt flashlight battery, and place them on opposite sides of your tongue, you will be able to "taste" the electricity as it flows across, producing a sour metallic flavor. What current is flowing through the saliva that coats your tongue?

__________ A

3.4 A certain lightbulb filament when hot has a resistance of 205 Ω. The potential difference across the filament at a certain instant is standard household voltage, 120 V. (a) What current is flowing through the bulb at that time? (b) If the filament of the bulb is a tungsten wire having a radius of 475 μm, what is the magnitude of the current density in the filament?

(a) __________ A
(b) __________ A/m2

3.5 You find a used battery and discover that it is no longer strong enough to power your CD player. Out of curiosity, you decide to find out what potential difference it actually supplies. You connect it to a 1.25 kΩ resistor and an ammeter (a device that measures current) and discover that the current through the circuit is 1.03 mA . What is the potential difference?

__________ V

Section 4 - Interactive problem: Ohm's law

4.1 Use the information given in the interactive problem in this section to calculate the resistance required to achieve the desired current and give Joe Pasty a nice tan. Test your answer using the simulation.

__________ Ω

Section 5 - Resistivity

5.1 An automotive junkyard crushes old vehicles into cubes 0.40 m on a side. Assume that these cubes are made of iron. Find the resistance you would expect to encounter if you connected one wire to the center of a face and one wire to the center of the opposite face, so that the two faces acted like the ends of a resistor.

__________ Ω

5.2 A length of nichrome wire (composed of nickel and chromium) is 75.0 cm long and has a rectangular cross section measuring 0.700 mm by 0.500 mm. The resistivity of nichrome is $1.08\times10^{-6}\ \Omega$m. If a potential difference of 225 V is applied between the two ends of the wire, how much current flows through it?

__________ A

5.3 You have a fine wire with a cross sectional area of 9.4×10^{-10} m^2 and a length of 1.1 m. You wish to determine what it is made out of, so you connect it to a standard AA battery (1.5 V) and measure a current of 0.048 A. What material is the wire most likely made out of? (Refer to the table in the text.)

i. Silver
ii. Copper
iii. Aluminum
iv. Iron

5.4 12-gauge copper wire has a radius of 1.03×10^{-3} m. (a) What is the resistance of 11.5 m of this wire at 20°C? (b) How much current will pass through the wire if a potential difference of 9.00 V is applied across it?

(a) __________ Ω
(b) __________ A

5.5 A section of wire 10.0 m long has a resistance of 4.00 Ω. (a) What is the resistance of 25.0 m of this wire? (b) What is the resistance of a 10.0 m section of wire of the same material with half the diameter? (c) What is the resistance of a wire whose length is 20.0 m and diameter is twice as large as the wire in part a?

(a) __________ Ω
(b) __________ Ω
(c) __________ Ω

5.6 The graphite in a pencil has a resistivity of 7.84×10^{-6} Ωm. The graphite's diameter is 0.700 mm and its length is 40.0 mm. (a) Find the resistance of this graphite "wire." (b) Find the diameter of a cylindrical copper wire of the same length (having resistivity 1.70×10^{-8} Ωm) that would produce the same resistance.

(a) ________ Ω
(b) ________ m

5.7 Silver has a resistivity of 1.62×10^{-8} Ωm, and a density of 1.05×10^{4} kg/m^3. (a) What is the volume of 7.50 g of silver? (b) You wish to fashion the 7.50 g of silver into a cylinder of length *L* and cross sectional area *A* so that the resistance between the ends of the cylinder is 3.50 Ω. What is *A*? (c) What is *L*? (Hint: Write two equations involving *L* and *A*, one for the resistance and one for the volume of the cylinder. Solve for *L* and *A*.)

(a) ________ m^3
(b) ________ m^2
(c) ________ m

Section 7 - Resistivity and temperature

7.1 A tungsten wire, initially at room temperature (20.0°C) with a resistance between its ends of 41 Ω, is placed in an oven. What will the temperature of the wire be when the resistance is 110 Ω? The temperature coefficient of resistivity of tungsten is 0.0050°C^{-1} at 20.0°C.

________ ° C

7.2 A thermometer can be made from a brass wire (temperature coefficient of resistivity 0.00150°C^{-1} at 20.0°C) which has a potential difference of 1.50 V between its ends and is connected to an ammeter. A measurement determines that the current flowing through the wire is 347 mA when the wire is at room temperature (20.0°C). (a) What is the resistance of the wire? (b) What change in current will accompany a decrease in temperature of one degree Celsius (or one kelvin)? Be careful of the sign of your answer. (c) How much current flows through the wire when the temperature is 35.0°C?

(a) ________ Ω
(b) ________ A
(c) ________ A

7.3 You have a resistor made of an unknown material. You note that at a temperature of 20.0°C its resistance is 529 Ω. Upon lowering its temperature to 0°C, the resistance drops to 476 Ω. (a) Find the temperature coefficient of resistivity of the material at 20.0°C. (b) Use the table in the text to make a guess as to the material the resistor is made of.

(a) ________ ° C^{-1}
(b)
i. Silver
ii. Carbon
iii. Tungsten
iv. Germanium
v. Silicon

7.4 A piece of iron has a resistance of 103 Ω at 40.0°C. Iron has a temperature coefficient of resistivity of 0.00640°C^{-1} at 20.0°C. What is the resistance of the iron at 20.0°C?

________ Ω

7.5 At 20.0°C, a mainly silicon resistor has a resistance of 585 Ω, and a tungsten resistor has a resistance of 752 Ω. At this temperature, the silicon resistor has a temperature coefficient of resistivity of −0.070°C^{-1} and the tungsten resistor has a coefficient of 0.0050°C^{-1}. At what temperature will their resistances be equal?

________ ° C

7.6 A circuit with silver wiring operates at 20.0°C. If you were to replace the silver with copper wiring of the exact same dimensions, at what temperature would you have to run the circuit in order to have the resistance of the copper circuit be the same as that of the silver circuit? (Use the resistivities measured at 20.0°C found in the table in the text.)

________ ° C

7.7 A conductor of an unknown substance has a resistance of 43.5 Ω at 40.0° C, and a resistance of 55.4 Ω at 50.0° C. What is the resistance of the conductor at 47.0° C? Hint: When applying the equation in the text, the temperature coefficient of resistivity must be measured at the temperature T_1.

________ Ω

Section 8 - Electric power

8.1 Instead of using batteries, a certain type of flashlight has a built in hand pump which you can squeeze to generate the power necessary to light the bulb. Suppose that squeezing the pump at a rate of twice per second generates 16 watts. If the light bulb has a resistance of 0.60 ohms, and you squeeze twice per second, what will be the potential difference across the light bulb?

__________ V

8.2 Many 120 V household circuits are rated at 15 A. This means that, if the current in the circuit exceeds this amount, there is a danger of overheated wires causing a fire. Such circuits are protected by circuit breakers that trip or fuses that blow and interrupt the current if it exceeds 15 A. (a) Is it safe to plug a 1400 W space heater into such a circuit? (b) Is it safe to plug two such space heaters into the same circuit, so that the power requirement is 2800 W?

(a) ○ Yes ○ No
(b) ○ Yes ○ No

8.3 How many 60 W light bulbs can you use in a 10 A, 120 V household circuit before tripping the circuit breaker and attempting to finish this homework problem in total darkness? The number of bulbs, times the individual power, equals the total power for the circuit.

__________ light bulbs

8.4 A certain appliance running at 240 V consumes 325 W of power. How much energy does it use in one hour?

__________ J

8.5 A simple heater for a cup of water consists of an insulated resistive coil that plugs directly into a 120 V electrical outlet. The heater is placed into a coffee mug that contains 0.550 kg of water and heats it to boiling (100°C) from room temperature (20°C) in 180 s. If the specific heat of the water is 4200 J/kg·K over this temperature range, what is the resistance of the coil? Ignore any changes in the resistance of the coil due to the increase in temperature.

__________ Ω

8.6 In a graphing calculator, there are four AAA batteries, which together provide a potential difference of 1.5 V and a total of 6.0 W·h of energy. If the batteries last 73 hours before they need to be replaced, what was the average current through the calculator?

__________ A

8.7 A 12 V car battery powers a light bulb. The bulb consumes 60 watts of power. (a) Assuming that the bulb is the only significant source of resistance in the circuit, find its resistance. (b) How much power would be consumed by a light bulb whose resistance was 1.0 Ω less?

(a) __________ Ω
(b) __________ W

Section 9 - Sample problem: solar panels

9.1 The power output of the Sun is 3.91×10^{26} W. At an energy cost of \$0.11/kW·h, what is the value of the energy produced by the Sun in one minute?

\$ __________

9.2 The intensity of sunlight at the distance of the Earth's orbit is 1380 W/m^2. An Earth-orbiting satellite has a solar panel that measures 1.35 m by 4.86 m, which converts solar energy to electrical energy with an efficiency of 26%. In one hour, how much electrical energy does the panel produce? Assume that the satellite's attitude control jets keep the panel oriented perpendicular to the incoming sunlight.

__________ J

Section 13 - Sample problem: power transmission

13.1 Power transmission lines often use a form of electric current called alternating current, but in many regions, such as the Province of Quebec, high-voltage direct-current lines are used instead. Direct current is the kind of electric current you are studying in this chapter. A certain direct-current power transmission line has a resistance of 0.255 Ω/km. 812 kV of potential drives the current from the generating station to a city located 125 km from the plant. What is the power loss due to resistance in the line?

__________ W

28.0 - Introduction

Capacitors store charge, and in doing so, they store energy. They are used to store energy in devices ranging from camera flashes to defibrillators, the systems used to “shock” the human heart back into its proper rhythm.

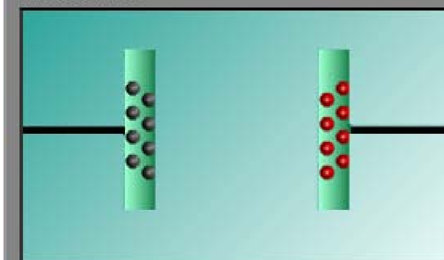
Capacitors are employed for other purposes as well. Because it takes a fixed amount of time to charge or discharge a given capacitor, capacitors are used in circuits where timing is essential, such as the tuner of a radio.

The "capacitor tree" (left) at the Fermilab National Accelerator Laboratory stored huge amounts of energy for high-energy accelerator experiments.

At the right, you see the design of a basic capacitor. It consists of two conducting plates separated by air. There is a wire connected to each plate, and these wires are attached to a source of potential difference that has caused the plates to become charged. The amounts of charge on the plates are equal in magnitude but opposite in sign.

You can launch the simulation and try an experiment that demonstrates one of the capacitor’s fundamental properties. In the simulation, you can change the potential difference across the plates with the ΔV controller next to the capacitor. When you change the potential difference, note what happens to the amount of charge stored by the capacitor. The approximate charge is displayed visually on the capacitor plates; for a more precise value you can rely on the charge readout gauge displayed below. Is the charge proportional to the potential difference?

interactive 1



Capacitor
Potential difference and charge

28.1 - Capacitors

Capacitor: A device with two conducting plates that can hold equal but opposite amounts of charge.

Capacitance: The ratio of the charge on one of the capacitor’s plates to the potential difference between the plates.

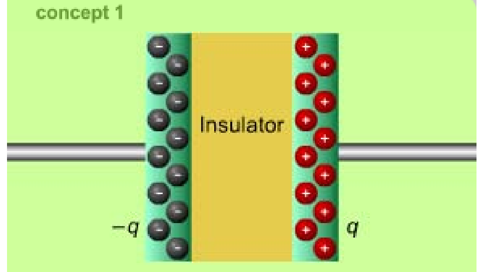
Various capacitors

The simplest capacitors consist of two parallel metal plates separated by a narrow gap. We use this configuration, a parallel-plate capacitor, for the definitions above. A battery (or other source of potential difference) causes the plates to become electrically charged: The plates contain equal but opposite amounts of charge. An insulator, which can be as simple as an air gap, separates the plates.

The capacitance of a capacitor tells how much charge it can store for a given potential difference between the plates; specifically, it equals the value of the positive charge on one plate in coulombs, divided by the potential difference in volts, a relationship shown in Equation 1. The amount of charge on **one** plate is represented by the letter q. The potential difference is also measured as a positive value, so charge, potential difference and capacitance are all positive values.

The geometry of a capacitor and the nature of its insulator are the two factors that determine its capacitance. Capacitors with larger plates, or with plates separated by a narrower distance, have greater capacitance.

concept 1



Capacitors
Composed of two conductors
Separated by insulator
Equal, opposite charges on conductors

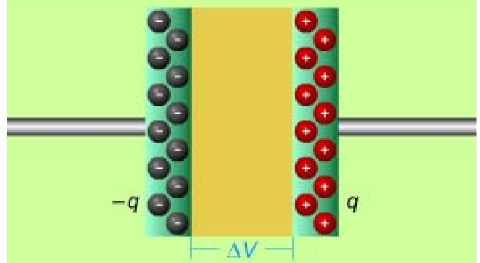
The other factor that can affect capacitance is the nature of the insulator between the plates. This insulator is called a *dielectric*. Air is one dielectric, but other dielectrics can be used to increase the capacitance.

The unit of measure of capacitance is the farad (F), named after the physicist Michael Faraday. A one-farad capacitor has a great deal of capacitance. It can store a lot of charge. Many commonly encountered capacitors have capacitances ranging from microfarads (10^{-6} F) down to picofarads (10^{-12} F). However, capacitors with greater capacitances certainly exist: car audio enthusiasts often boast systems employing one-farad capacitors.

The movement of electrons creates the charges on the two plates of a capacitor. They move away from one plate, leaving a net positive charge behind, and go toward the other plate, which takes on a negative charge. Note that the electrons do **not** move across the gap between the plates. Rather, they move along the wires whose ends you see in the illustrations.

A source of potential difference such as a battery exerts a force on the electrons, causing them to move. This force through a distance constitutes work. The amount of work done during the charging process equals the electric potential energy stored by the plates. The greater the charge on its plates, the greater the amount of energy a particular capacitor is storing.

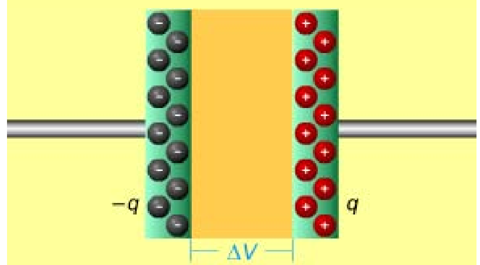
concept 2



Capacitance

Relationship between charge, potential difference

equation 1



Capacitance

$$C = \frac{q}{\Delta V}$$

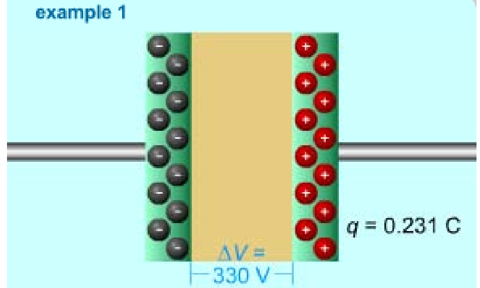
C = capacitance

q = positive charge on one plate

ΔV = potential difference across plates

Units: farads (F)

example 1



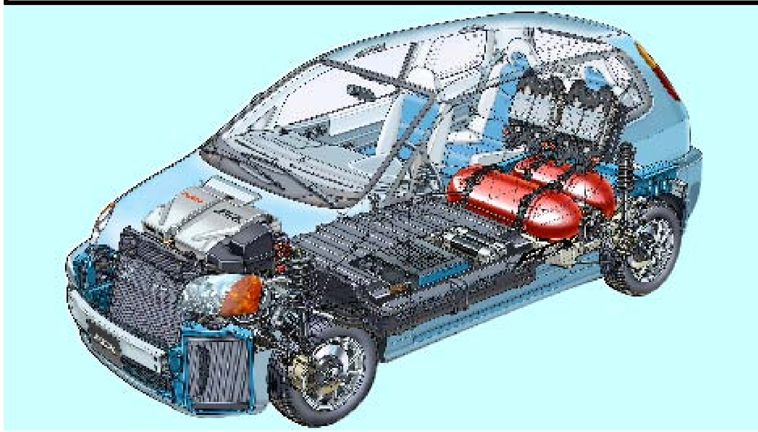
What is the capacitance of this capacitor?

$C = q/\Delta V$

$C = 0.231 \text{ C} / 330 \text{ V}$

$C = 0.000700 \text{ F} = 700\ \mu\text{F}$

28.2 - Interactive checkpoint: the fuel-cell car



How much charge does an 8.00-farad ultracapacitor accumulate when it is charged with a potential difference of 200 V?

Honda's FCX fuel-cell automobile uses a custom built, 8.00-farad *ultracapacitor* that provides power in faster bursts than a fuel cell can deliver on its own, allowing the car to accelerate quickly.

The ultracapacitor is charged by both the car's fuel cell and energy recovered each time the car brakes. Together, these energy sources can be used to provide a potential difference of 200 V across the capacitor electrodes (plates).

Answer:

$q =$ ______ C

28.3 - Gauss' law and the capacitance of a parallel-plate capacitor

In this section we derive the formula for the capacitance of an ideal parallel-plate capacitor in terms of the capacitor's geometry: its area and the distance between the plates. The illustration for Equation 1 is a conceptual diagram of a parallel-plate capacitor. We assume that a vacuum separates the plates.

The formula shown on the right provides a good approximation of the capacitance of the capacitor. The smaller the distance between the plates of an actual capacitor, relative to their length and width, the more accurately this equation describes capacitance. In the derivation, we treat each parallel plate as an infinite plane of charge, and focus our attention on the uniform field between the centers of the plates, ignoring the more complex field at their edges.

To derive an equation for capacitance, we need to use equations that describe the charge on the plates and the potential difference between them. We do this by analyzing the field generated by the plates.

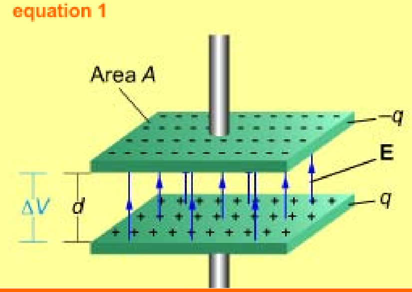
Variables

capacitance of capacitor	C
charge on one plate	q
potential difference	ΔV
area of one plate	A
distance between plates	d
total flux through Gaussian surface	Φ_E
permittivity constant	$\varepsilon_0 = 8.85\times10^{-12}\ \mathrm{C^2/N{\cdot}m^2}$
electric field between plates	$\mathbf{E}$
field due to positive plate	$\mathbf{E}_{pos}$

Strategy

1. First we relate charge and electric field using Gauss's law. The diagram of Equation 2 shows the Gaussian surface we use for the field due to the positive plate. (This parallels the work done in deriving the equation for the strength of the electric field due to an infinite plane of positive charge. Consider it extra practice using Gauss' law, if you like.) By relating the field to the amount of charge per unit area of one capacitor plate, and knowing the area of the plates, we quantify the **charge** necessary to determine the capacitance.

equation 1

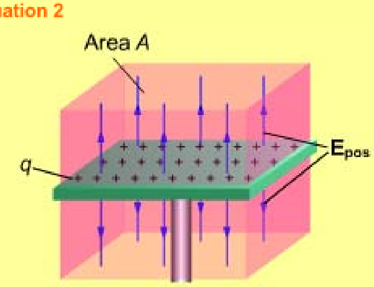


Parallel-plate capacitor

$$C = \frac{\varepsilon_0 A}{d}$$

C = capacitance
ε_0 = permittivity constant
A = area of plate
d = distance between plates

equation 2



Derivation of capacitance formula

2. We also need an expression that tells the other quantity in the definition of capacitance, the **potential difference** between the plates, in terms of their geometry. We obtain this equation by using the relationship of the potential difference and distance in a uniform field: $\Delta V = Ed$.
3. Then, we use the definition of **capacitance** to combine the formulas generated by the previous strategy steps. By combining the expressions we have developed, we generate an equation that describes the capacitance in terms of the geometry of the capacitor.

Physics principles and equations

An infinite charged plane having uniform charge density produces a uniform electric field that is perpendicular to the plane on either side. For a uniform electric field perpendicular to a surface, the electric flux equals the field strength times the surface area.

Electric fields can be added like vectors (the principle of superposition).

In a uniform electric field, the field strength equals the positive potential difference between any two points divided by the distance d parallel to the field between them. In this case, d is the distance between the capacitor plates.

$$E = \Delta V/d$$

We also use Gauss' law:

$$q = \varepsilon_0 \Phi_E$$

The definition of capacitance is

$$C = q/\Delta V$$

Step-by-step derivation

We start the derivation by applying Gauss' law to the field $\mathbf{E}_{pos}$ generated by the positive plate of the capacitor. The five steps that follow are a terse restatement of work already done in the study of electric flux when we applied Gauss' law to an infinite plane of charge. Ignoring the field variations at the edge of the plate, we can consider the field generated by a plate to be uniform.

Step	Reason
1. $\Phi_E = 2E_{pos}A$	flux for Gaussian surface
2. $\Phi_E = q/\varepsilon_0$	Gauss' law
3. $E_{pos} = q/2\varepsilon_0 A$	combine equations 1 and 2 and solve for E_{pos}
4. $E = q/\varepsilon_0 A$	net field from two plates
5. $q = \varepsilon_0 AE$	solve for q

The field strength of the uniform field $\mathbf{E}$ between the plates equals the positive potential difference between any two points in the field divided by the distance between them. For this relationship to hold true, the displacement must be against the field, and it is: the magnitude of the displacement we care about is d, the distance between the plates.

Step	Reason
6. $E = \Delta V/d$	equation stated above
7. $\Delta V = Ed$	solve for ΔV

In our final steps, we apply the definition of capacitance. We use the formulas derived in steps 5 and 7 to determine the expression for the capacitance in terms of the geometric attributes of the parallel-plate capacitor.

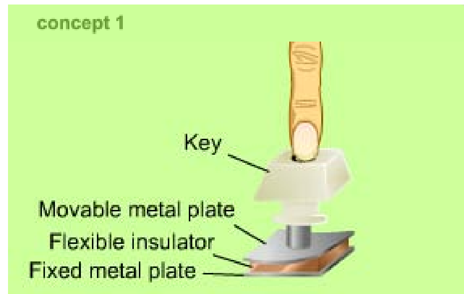
Step	Reason
8. $C = \dfrac{q}{\Delta V} = \dfrac{\varepsilon_0 AE}{Ed}$	substitute equations 5 and 7 into definition of capacitance
9. $C = \dfrac{\varepsilon_0 A}{d}$	simplify

The final step yields an equation for the capacitance of a parallel-plate capacitor in terms of its physical configuration. The capacitance increases with the area of the plates and decreases with the distance between them.

28.4 - Physics at work: capacitors and computer keyboards

Your computer keyboard may contain capacitors. In some types of keyboards, each key sits above its own capacitor. Although more expensive to produce than those based on other designs, capacitor based keyboards are more reliable. The bottom part of the capacitor is fixed, and the top is a movable plate. A springy insulating material separates the plates. As you press down on a key, the two plates of its capacitor are pushed closer together. Since this changes the geometry of the capacitor, it also changes the capacitance: Specifically, it increases it. A microprocessor in the keyboard interprets the change in capacitance as a signal that the key has been pressed.

The keyboard microprocessor relays this information to the computer's central processing unit, which responds to it according to which application is running. For example, a word processor might display the keystroke as a character, while a game might interpret it as a command for a player to shoot a basketball.



Computer keyboards
Keystroke reduces distance between plates
Microprocessor detects changed capacitance

28.5 - Interactive problem: capacitor plate separation

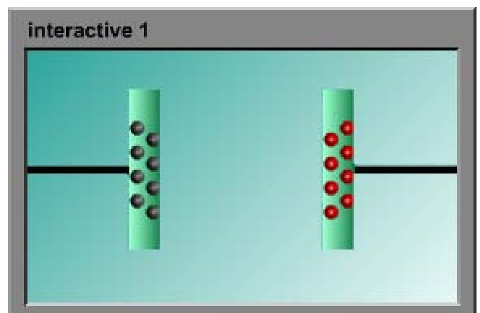
You have seen how the charge that can be stored by a capacitor for a given potential difference is measured by its capacitance. For an ideal parallel-plate capacitor, the capacitance can be calculated from the area and separation of the two plates.

At the right, you see a parallel-plate capacitor, similar to the one in the introduction to this chapter. The plates will have a potential difference of 175 volts across them when you press GO to close a switch.

Your mission is to set the distance (in millimeters) between the plates so that the charge on one plate is 2.00 nC. The plates are squares, each exactly 10 cm on a side, so the surface area of each plate is 0.0100 m^2. Be careful with these varied units!

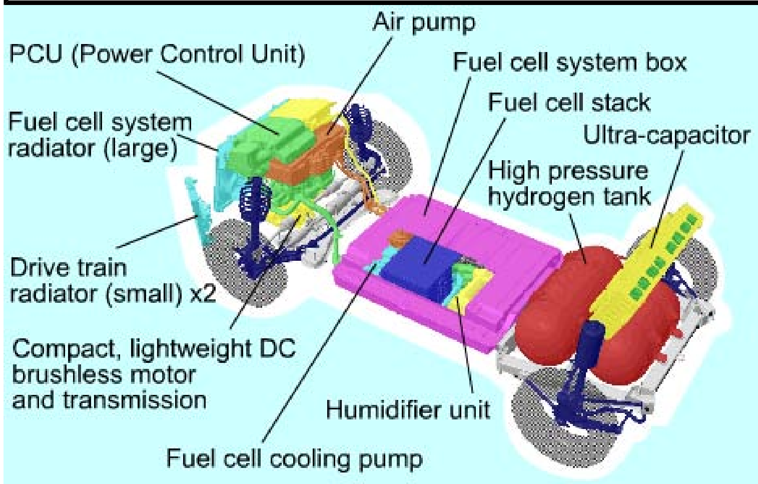
Once you have made your calculations and set the distance, press the GO button. It closes a switch and the charge will flow. (The charge flows more slowly than it would in reality to add a little suspense...)

If you have any trouble finding the correct distance, you may wish to review earlier sections of this chapter, which introduced the concept of capacitance and related it to the dimensions of a parallel-plate capacitor.



Capacitor configuration
Adjust separation for desired charge

28.6 - Interactive checkpoint: engineering the fuel-cell car



The ultracapacitor in the Honda FCX fuel-cell car has a huge capacitance of 8.00 F, yet it is still small enough to fit right behind the back seat of the car.
If you were to construct a parallel-plate capacitor with two plates separated by a 1.00 mm air gap, what area would the plates need in order to have the same capacitance as Honda's ultracapacitor? (Hint: It would not fit behind the seat!)

Answer:

A = [ ] m^2

The answer indicates that a simple, square parallel-plate capacitor would have to be enormous to have the desired capacitance, more than 30 km on a side! Something that would hardly fit behind the back seat, even if this were an SUV. Even if, through a superhuman feat of engineering, you succeeded in reducing the plate separation to 10 μm (about one-tenth the width of a human hair) the capacitor would still have to exceed 3 km on a side.

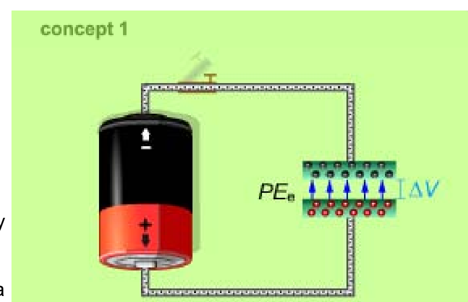
Clearly the ultracapacitor is not an ideal parallel-plate capacitor. The ultracapacitor utilizes special microscopically porous electrodes that pack an extremely large surface area into a very small volume.

28.7 - Energy in capacitors

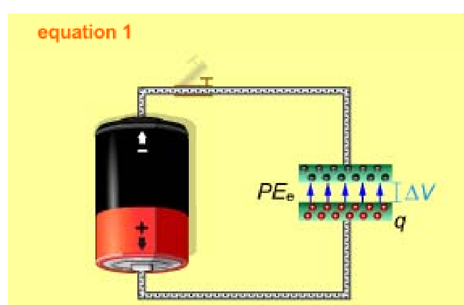
Capacitors store energy. A battery or other device performs work as its electric field pulls electrons from the positive plate of a capacitor and pushes them to the negative plate. The force of this field drives electrons away from the attractive force of the positive plate and toward the repulsive force of other electrons already on the negative plate. This work increases the electric potential energy of the capacitor.

To the right you see two equations that express the energy stored in a capacitor. The first equation is stated in terms of the charge on the capacitor; the second is stated in terms of the potential difference across the capacitor. We can derive these equations by calculating the work required to move each charge from one plate to the other.

The example problem asks you to find the amount of energy stored in a capacitor with a potential difference of 25,000 V across its plates. This is representative of the voltages found in television sets. The stored energy in a TV capacitor can give you a dangerous shock, even after the set is unplugged. This is why the back panels of televisions are labeled with warnings against attempting to disassemble them yourself. We recommend that you follow this advice!



Energy in a capacitor
Battery's field causes electrons to move
Charges create electric field
Electric field stores energy



Energy in a capacitor

$$PE_e = \frac{q^2}{2C}$$

$$PE_e = \frac{C(\Delta V)^2}{2}$$

PE_e = electric potential energy
q = charge on capacitor plate
C = capacitance
ΔV = potential difference

What is the electric potential energy stored in this capacitor?

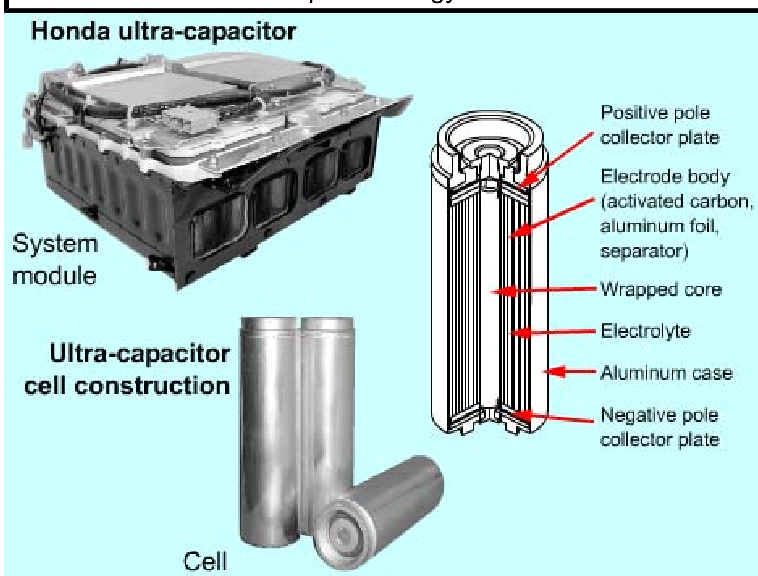
$PE_e = C(\Delta V)^2 /2$

$PE_e = (0.0015\ \mu\text{F})(25{,}000\ \text{V})^2 /2$

$PE_e = (1.5\times10^{-9}\ \text{F})(2.5\times10^{4}\ \text{V})^2 /2$

$PE_e = 0.47\ \text{J}$

28.8 - Interactive checkpoint: energy in a fuel-cell car



An FCX fuel-cell car brakes and stops at a traffic signal. During this time, its 8.00-farad ultracapacitor gets fully charged by a combined potential difference of 200 V generated by the brakes and the fuel cell. When the light turns green, the 1680-kilogram car initially uses only energy from the capacitor to accelerate.

How fast will the car be moving when the capacitor is fully discharged?

In this problem, assume that road friction, air resistance, and electrical resistance use up 20.0% of the energy available, and that the rest of the electric potential energy stored in the capacitor's electric field is transformed into the kinetic energy of the car.

State your answer in miles per hour (1.00 m/s = 2.24 mph).

Answer:

v = ______ mph

28.9 - Interactive problem: You are Zeus

In honor of the XXVIII Olympiad in Athens, this interactive exercise takes you back to the very first Olympiad in ancient Greece. It is the year 776 BCE. You play Zeus, and your job is to use a lightning bolt to light the Olympic flame. In addition to injecting mythology, we significantly simplify some atmospheric science. But we think the result − some fun with capacitors − is worth it.

The source of lightning bolts can be modeled as a huge capacitor: regions of positive and negative charge in adjacent cloud layers. The exact process through which these charged regions arise is complex and subject to debate, but standard theories suggest that it is akin to common sources of everyday static electricity, atoms and molecules brushing by each other at the edges of powerful updrafts and downdrafts of air. Excess electrons congregate in one region, leaving a net positive charge in another. When the charge separation becomes great enough, electricity discharges between the clouds, or from a low hanging cloud to the ground (lightning!).

interactive 1

First Olympiad
Light the Olympic flame

In this simulation, you are Zeus if you like, and you want to create a charge distribution that stores 255 million joules of energy. The cloud layers here form a capacitor of atypical proportions: two parallel plates, each with an interior surface area of 1.00 km^2 separated by a distance of 177 meters. You are asked to treat this as an ideal parallel-plate capacitor, and to analyze it as though there were a vacuum between the "plates."

Start by determining the required capacitance. The geometrical specifications may help you, or you can use the fact that, with the initial potential difference of 5.60×10^7 V across the clouds, the required charge on one "plate" is 2.80 C.

As mentioned above, you want the clouds to contain 255 million joules of energy. Using your divine powers, you can increase the potential difference between the clouds. Set the correct potential difference in the simulation and then press the ZAP button. If you are correct, you will know!

If you need to review, see the earlier section on energy in capacitors.

28.10 - Capacitors and electric field energy density

Electric field energy density (u_E): The amount of electric potential energy stored in an electric field per unit volume.

In this section, we discuss energy density in the context of electric fields. We then derive a formula for the energy density of an electric field, using a parallel-plate capacitor as the source of the field.

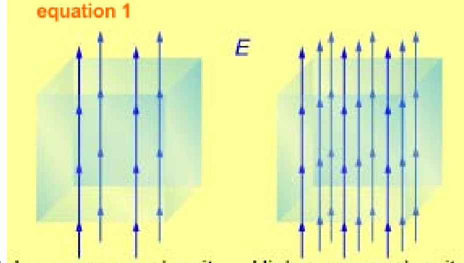
The electric field energy density u_E equals the amount of electric potential energy stored in an electric field per unit volume. You see this definition expressed in Equation 1 to the right. In the definition we write the volume as Vol to avoid any possible confusion with the symbol for potential difference, ΔV.

We can calculate the energy density of an electric field using the capacitor configuration shown in Equation 2. The resulting equation, which is derived here, states that the energy density is proportional to the square of the electric field strength. Although this equation is derived using a parallel-plate capacitor, it holds true for any electric field from any source.

Variables

electric potential energy	PE_e
capacitance of capacitor	C
potential difference between plates	ΔV
area of one plate	A
distance between plates	d
volume between plates	Vol
permittivity constant	$\varepsilon_0 = 8.85 \times 10^{-12}\ \mathrm{C^2/N \cdot m^2}$
electric field strength	E
energy density of electric field	u_E

equation 1

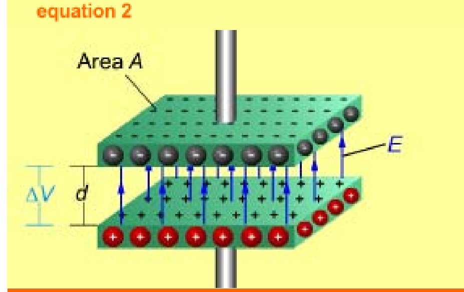


Electric field energy density

$$u_E = PE_e / Vol$$

u_E = electric field energy density
PE_e = electric potential energy
Vol = volume

equation 2



Energy density formula

Strategy

1. Begin by stating three equations developed in other sections.
2. Combine these equations to write a formula for the potential energy stored by the capacitor in terms of the geometry of the capacitor and the electric field strength. Also, write a formula for the volume of the region between the plates of the capacitor.
3. Finally, substitute the expressions for potential energy and volume just derived into the definition of energy density to obtain a formula for the electric field energy density in terms of field strength.

$$u_E = \tfrac{1}{2}\varepsilon_0 E^2$$

ε_0 = permittivity constant
E = electric field strength

Physics principles and equations

The capacitance of a parallel-plate capacitor is

$$C = \varepsilon_0 A/d$$

The field between the plates of an ideal parallel-plate capacitor is uniform. The following equation relates potential difference to a distance traveled against a field line in a uniform field.

$$\Delta V = Ed$$

The electric potential energy stored in a capacitor is

$$PE_e = C(\Delta V)^2/2$$

Step-by-step derivation

We start by restating three equations.

Step	Reason
1. $C = \varepsilon_0 A/d$	capacitance of parallel-plate capacitor
2. $\Delta V = Ed$	equation stated above
3. $PE_e = C(\Delta V)^2/2$	energy stored in capacitor

Now we will derive the formulas for the electric potential energy and volume in terms of the geometry of the capacitor and the electric field strength.

Step	Reason
4. $PE_e = \left(\dfrac{\varepsilon_0 A}{d}\right)E^2 d^2/2$	substitute equations 1 and 2 into equation 3
5. $PE_e = \varepsilon_0 A d E^2/2$	simplify
6. $Vol = Ad$	by inspection

In the following steps we expand the numerator and denominator of the definition of electric field energy density, using expressions found in the previous stage of the derivation.

Step	Reason
7. $u_E = PE_e/Vol$	definition
8. $u_E = \dfrac{\varepsilon_0 A d E^2/2}{Ad}$	substitute equations 5 and 6 into equation 7
9. $u_E = \tfrac{1}{2}\varepsilon_0 E^2$	simplify

28.11 - Physics in medicine: defibrillator

Although electricity can be harmful if misused, doctors have used it for medicinal purposes since the days of the ancient Greeks. For example, early physicians used discharges from the electric torpedo fish to relieve aches and pains.

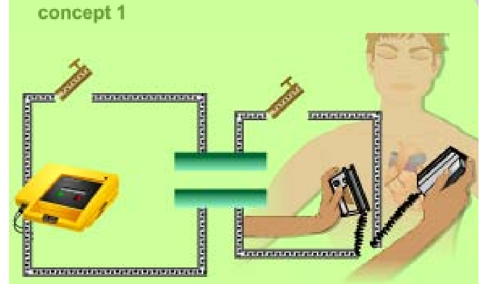
Electricity continues to be used in medicine. If you have ever watched the television show *ER*, chances are you have become well acquainted with a device called a *defibrillator*. The heart relies on electrical impulses to cause its muscles to contract. When the muscle cells of the heart begin to contract out of synchronization, or to fibrillate, a defibrillator can be used to jolt them back into a synchronous rhythm. Contrary to popular belief, defibrillators are not used to "restart" a heart that has stopped beating. You may get the impression from TV shows that doctors use defibrillators as frequently as stethoscopes, but that is just one of the exaggerations of television. However, these instruments are becoming standard equipment in ambulances and on airplanes.

A conceptual diagram of a defibrillator is shown on the right. The source of potential difference on the left charges the capacitor. When the switch on the right is closed, the capacitor rapidly discharges, sending an electrical current through the patient's heart.

The example problem shows how to calculate the average current when the discharge time, capacitance and potential difference are known. The current is large, but it lasts for just a few milliseconds, providing a powerful but brief shock.

Cardiac defibrillator.

concept 1

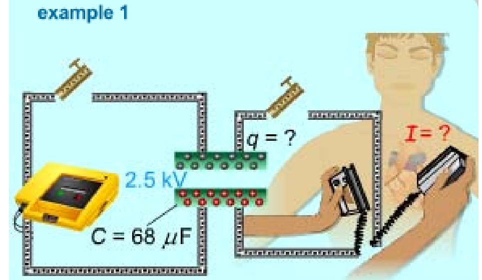


Defibrillator

Capacitor used in heart defibrillator

- Battery charges capacitor to 2500 V
- Discharges in milliseconds, sending large current through heart

example 1



The defibrillator capacitor discharges in 11 ms. How much current does it send through the heart?

$C = q/\Delta V$

$q = C\Delta V = (68 \times 10^{-6}\ \text{F})(2.5 \times 10^{3}\ \text{V})$

$q = 0.17\ \text{C}$

$I = \Delta q/\Delta t$

$I = (0.17\ \text{C}) / (0.011\ \text{s})$

$I = 15\ \text{A}$

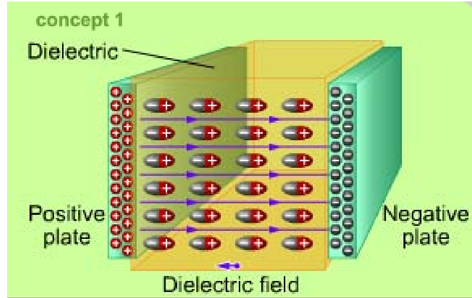
28.12 - An insulator (dielectric) in an electric field

The simplest capacitor is made up of two plates separated by a vacuum. In this section we will discuss what happens when an insulating material, a *dielectric*, is placed between the plates.

Like a vacuum, this insulating material is intended to prevent the flow of charge between the plates, but it has another effect, as well: It reduces the overall strength of the field between the plates, and this proves to be a desirable effect.

To understand how a dielectric reduces the net field, let's first consider what happens when an atom is placed in the electric field generated by two charged plates. It might seem that nothing would happen because the atom is electrically neutral.

However, an atom is made up of a positive nucleus surrounded by negative electrons. If the field is very strong, it can *ionize* atoms, separating electrons from the atoms. At this point, the dielectric material becomes a conductor: Current can flow between the plates, causing the capacitor to break down. Although an interesting phenomenon to witness, this outcome is not relevant to the rest of this section.



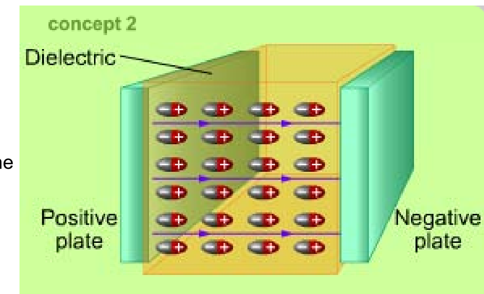
Dielectric
When plates are charged:
Dipole alignment creates field in dielectric
· Dielectric field opposes field of plates

Less extreme fields can *polarize* an atom, in essence stretching it so that its electrons tend to be on one side and its positive nucleus is on the other. This turns the atom into a dipole, a body with positively and negatively charged regions.

Some kinds of molecules, such as water molecules, are always dipoles. This type of molecule (called a *polar molecule*) has regions of positive and negative charge based on its structure. When no external electric field is present, the dipoles of substances such as water are randomly aligned, and they create no net electric field.

When a substance containing dipoles is placed in the electric field between a charged capacitor's plates, the field will cause some of the dipoles to align so that their positive poles point toward the negative plate and their negative poles point toward the positive plate.

Many, but by no means all, of the dipoles that compose the substance will align in this manner with the external field. Their random thermal motion provides a constant counterbalance to the organizing tendency of the field. The effect of the alignment that does occur is to weaken the overall field between the plates. There are two ways to understand the weakening of the capacitor field in the presence of a dielectric.



Dielectric effect
Dielectric weakens overall capacitor field

First, consider the net effect of the individual fields of all the aligned dipoles. Each aligned dipole has its own electric field, oriented in the direction opposite to the field created by the plates. This is shown in Concept 1. The aligned dipoles' fields reduce the overall field between the plates because they point oppositely to the capacitor's field.

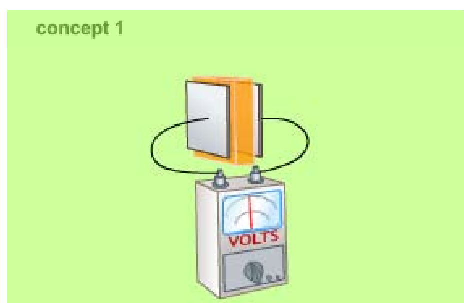
You can also look at the diagram in Concept 2 to understand why the overall field is diminished. In the bulk of the dielectric the net charge is zero. This is because every aligned positive pole there is next to an aligned negative pole on a neighboring dipole, which it balances out. However, on the far left there is an unbalanced layer of negative charge, and on the far right there is an unbalanced layer of positive charge. This means there is a layer of negative charges adjacent to the positive plate of the capacitor and a layer of positive charges adjacent to the negative plate.

For this reason the entire dielectric material can be said to be polarized (to have positive and negative regions). This polarization decreases the net charge in close proximity to the surface of each capacitor plate. The net charge of each plate is now the charge on the plate surface minus the charge on the adjacent face of the dielectric. Less effective charge on each plate surface means the overall field between the plates is weaker.

28.13 - Dielectrics

Dielectric: An insulating substance placed between the plates of a capacitor to increase its capacitance.

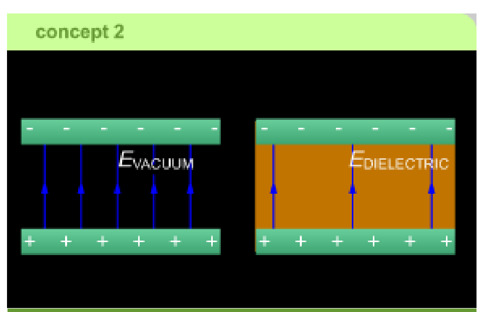
Dielectric constant (κ): Measures reduction in electric field caused by a dielectric. It equals the ratio of the field in a vacuum to the field in the dielectric.



Dielectric
Insulator that increases capacitance

Dielectrics are insulating materials used in capacitors to increase their capacitance. Effective dielectrics make possible the manufacture of small, high-farad capacitors.

The materials discussed in this section are *linear dielectrics*. The dielectric field of a linear dielectric is linearly proportional to the strength of the external electric field.



Dielectrics and electric fields

The dielectric constant, represented by κ (Greek letter kappa), is a property of a material. It equals the ratio of two electric fields: the electric field inside a capacitor with a vacuum separating the plates and the field strength for the same capacitor charge with the dielectric present. This is stated in Equation 1. The greater the dielectric field within a dielectric, the more it diminishes the field caused by the plates, and the greater the value of the dielectric constant.

Dielectrics increase the capacitance of capacitors. As is stated in Equation 2, the capacitance with the dielectric present equals the capacitance without it times the dielectric constant.

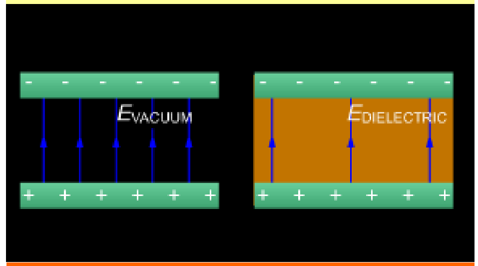
The table in Equation 3 lists some dielectric constants. As you can see, the dielectric constants for vacuum and air are quite close. Strontium titanate, a substance used in commercial capacitors, has a far greater dielectric constant, especially at low temperatures, than a vacuum or air.

Dielectrics also can be classified by their *dielectric strength*. The dielectric strength characterizes the field strength at which the dielectric becomes a conductor and charge will flow through it. This is not a desirable effect, so capacitors are often labeled with their maximum safe field strength (in V/m). Like many stress properties, the dielectric strength of a material can be difficult to determine precisely, so the stated dielectric strength should be an approximate, conservative value. The dielectric strengths of selected materials are also shown in the table.

Lightning is a very visible example of the breakdown of a dielectric − air − between oppositely charged objects, such as two clouds or a cloud and the nearby surface of the Earth. When the electric field between them becomes strong enough, the separating atmosphere becomes a conductor and current flows through it as a lightning bolt.

Dielectric diminishes electric field

equation 1

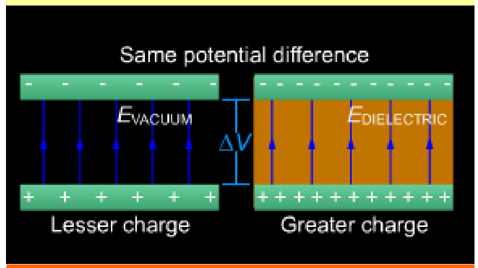


Definition of dielectric constant

$$\kappa = E_{\text{vacuum}}/E_{\text{dielectric}}$$

κ = dielectric constant
E = field in vacuum or dielectric

equation 2



Capacitance with dielectric present

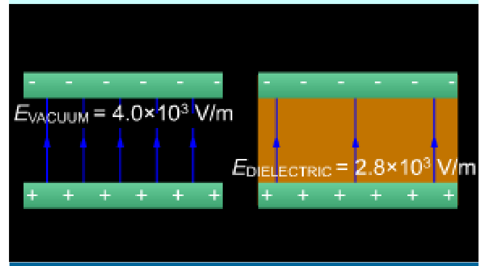
$$C_\kappa = \kappa C$$

C_κ = capacitance with dielectric
κ = dielectric constant
C = capacitance with vacuum

equation 3

	Dielectric constant	Dielectric strength (in 10^6 V/m)
Vacuum	1	n/a
Air	1.00054	3.0
Paper	1.7 to 2.6	4 to 9
Rubber	2 to 3.5	20 to 27
Glass	5.4 to 9.9	30 to 150
Water (293K)	80.20	65 to 70
Strontium titanate (298K, 78K)	332, 2080	

Table of dielectric constants



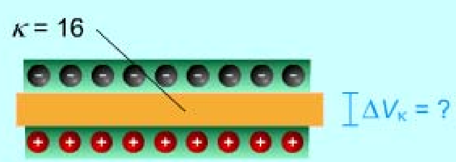
The capacitors have identical dimensions and charges. What is the dielectric constant of the dielectric on the right?

$\kappa = E_{\text{vacuum}}/E_{\text{dielectric}}$

$\kappa = (4.0 \times 10^3 \text{ V/m}) / (2.8 \times 10^3 \text{ V/m})$

$\kappa = 1.4$ (dimensionless)

28.14 - Sample problem: dielectric and potential difference



A battery causes the plates of a capacitor to charge so that the potential difference between them is 9.0 V. The capacitor is then isolated so that its charge remains constant and the battery no longer acts on it.

After a dielectric (K = 16) is inserted between the plates, what is the potential difference between them?

Variables

initial potential difference	$\Delta V = 9.0$ V
initial capacitance	C
charge on capacitor	q
dielectric constant of material	$\kappa = 16$
capacitance with dielectric present	C_κ
potential difference with dielectric present	ΔV_κ

Strategy

1. Use the definition of capacitance, and the equation for capacitance with a dielectric present, to determine the potential difference when the dielectric is present.

Physics principles and equations

The definition of capacitance is

$$C = q/\Delta V$$

The relation between capacitances with and without a dielectric present is

$$C_\kappa = \kappa C$$

Charge in any isolated system is conserved.

Step-by-step solution

We use the definition of capacitance twice; once when the dielectric is not present, and again when it is. The amount of charge remains constant since the capacitor is isolated.

Step	Reason
1. $q = C\Delta V$	definition of capacitance
2. $C_\kappa = \kappa C$	capacitance with dielectric
3. $\Delta V_\kappa = \dfrac{q}{C_\kappa}$	definition of capacitance

Now we combine the above expressions to find a relation between the potential differences with and without a dielectric present. We evaluate the resulting equation to answer the problem.

Step	Reason
4. $\Delta V_\kappa = \dfrac{C\Delta V}{\kappa C}$	substitute equations 1 and 2 into equation 3
5. $\Delta V_\kappa = \dfrac{\Delta V}{\kappa}$	simplify
6. $\Delta V_\kappa = \dfrac{9.0\text{ V}}{16} = 0.56\text{ V}$	evaluate

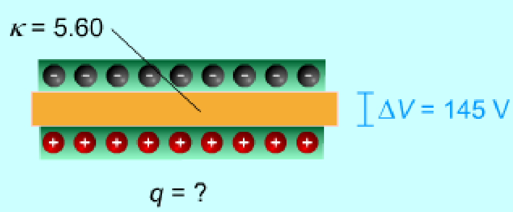
The potential difference between the plates after inserting the dielectric turns out to be the original potential difference divided by the dielectric constant. The higher the dielectric constant, the more the potential difference will be decreased after the dielectric is introduced.

This result is interesting when applied to the potential energy $PE_e = ½C(\Delta V)^2$ stored by a capacitor. After you insert a dielectric, the potential energy becomes

$$PE_\kappa = ½C_\kappa(\Delta V_\kappa)^2 = ½(\kappa C)\left(\frac{\Delta V}{\kappa}\right)^2 = \frac{½C(\Delta V)^2}{\kappa} = \frac{PE_e}{\kappa}$$

In other words, the potential energy decreases when you insert a dielectric between the plates of a capacitor. Why? The capacitor is doing work on the dielectric as you insert it, rotating many of its small dipoles into alignment.

28.15 - Interactive checkpoint: dielectric in a capacitor



A parallel plate capacitor has a capacitance of 3.50 μF when a vacuum separates its plates. You insert an insulator with a dielectric constant of κ = 5.60 and apply a potential difference of 145 V across the plates.

How much charge builds up on a plate?

Answer:

q = ______ C

28.16 - Physics at work: commercial capacitors

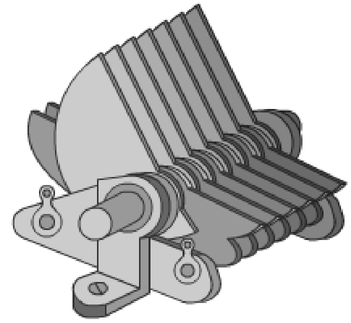
Capacitors in electronic equipment, like a stereo system, that you might have at home are often made from two layers of metallic foil backed by thin sheets of a dielectric material. When these layers are rolled up, the result is a multilayer cylindrical capacitor with alternating positive and negative cylindrical conductors, separated by the dielectric material. Several capacitors of this type are shown in Concept 1.

Higher voltage capacitors can consist of sets of thin parallel metallic plates immersed in insulating silicone oil. Other capacitors, capable of storing large amounts of charge, are made with insulators consisting of ceramic materials that can have very high dielectric constants.

A *variable capacitor*, like the one in the picture above, features interleaved metal plates. You change the capacitance by rotating the knob, increasing or decreasing how much of the area of each plate is close to the neighboring plates on either side. Devices like this can be used to change the capacitance of a circuit in a radio tuner, altering what is called the resonant frequency of the circuit and enabling it to “tune in” a radio station broadcasting at a particular frequency.

Electrolytic capacitors can operate with very high potential differences across their two conducting surfaces. An electrolyte is a nonmetallic conductor, or a substance that when dissolved in a suitable solvent becomes a conductor. In an electrolyte, current is composed of ions, not electrons.

In an electrolytic capacitor, one of the conductors is a solid or porous metal electrode, the dielectric is an insulating metal oxide coating on the surface of the electrode, and the other conductor is an electrolytic liquid or solid. Since the distance between the “plates” is the thickness of the insulating metal oxide, which is quite small, tiny electrolytic capacitors can have extremely high capacitance. The capacitors used in certain fuel-cell automobiles are a form of electrolytic capacitor.



A variable capacitor used for tuning a radio.

concept 1

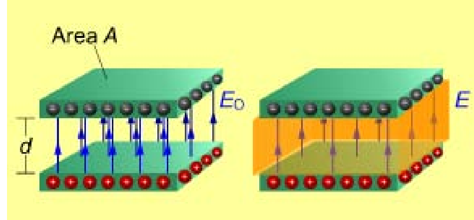
Commercial capacitors
Rolled foil capacitors

28.17 - Derivation: capacitance with a dielectric present

In this section, we derive an equation for the capacitance of a parallel plate capacitor with a dielectric present. Note that the parenthetical expression in the equation represents the capacitance of an ideal parallel-plate capacitor with a vacuum between its plates. This derivation proves a specific case of a previously stated (but not proved) equation that holds true for all types of capacitors: $C_\kappa = \kappa C$.

Two parallel-plate capacitors are shown to the right. Each capacitor has the same dimensions and the same charge. Their capacitances and the potential differences across them are not the same.

equation 1



Capacitance of a parallel-plate capacitor with dielectric

$$C_\kappa = \kappa\left(\frac{\varepsilon_0 A}{d}\right)$$

C_κ = capacitance w/ dielectric present
κ = dielectric constant of material
ε_0 = permittivity constant
A = area of plate
d = distance between plates

Variables

field strength in vacuum	E_0
field strength with dielectric	E
charge on one plate of capacitor	q
permittivity constant	$\varepsilon_0 = 8.85\times10^{-12}\ \mathrm{C^2/N{\cdot}m^2}$
area of each plate	A
potential difference with dielectric	ΔV_κ
distance between plates	d
dielectric constant of material	κ
capacitance with dielectric	C_κ

Strategy

1. Express the strengths of both the vacuum field E_0 and the reduced field E in terms of the properties of the capacitor, including the charge q and the area in one case, and the potential difference ΔV_κ between the plates in the other.
2. Relate the resulting expressions for E_0 and E, using the definition of the dielectric constant.
3. The previous steps result in an equation that contains both q and ΔV_κ. Use the definition of capacitance to replace these two quantities with the capacitance C_κ of the parallel-plate capacitor with the dielectric material present, and simplify to obtain the desired equation.

Physics principles and equations

We will use the equation for the strength of the electric field between parallel charged plates in a vacuum:

$$E_0 = \frac{q}{\varepsilon_0 A}$$

The field between the conductors in an ideal parallel-plate capacitor is uniform. The equation below holds true in a uniform field, where d is a distance measured parallel to the field lines:

$$\Delta V = Ed$$

The definition of the dielectric constant is

$$\kappa = E_0/E$$

The definition of capacitance is

$$C = q/\Delta V$$

Step-by-step derivation

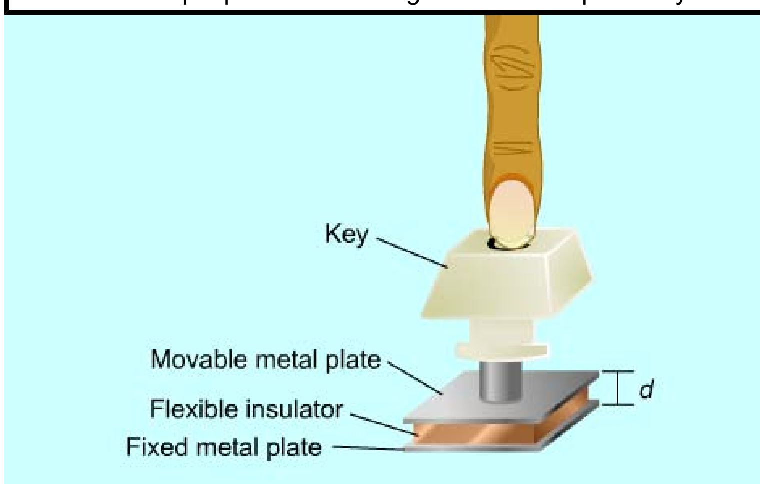
The definition of the dielectric constant relates the field strength in a vacuum to the field strength with the dielectric present. Our goal is to replace the field strengths with expressions more closely related to capacitance.

Step	Reason
1. $E_0 = q / \varepsilon_0 A$	field strength between plates in vacuum
2. $E = \Delta V_\kappa / d$	potential difference in uniform field
3. $\kappa = E_0 / E$	definition of dielectric constant
4. $\kappa = \dfrac{q/\varepsilon_0 A}{\Delta V_\kappa / d}$	substitute equations 1 and 2 into equation 3

The remainder of the work involves algebraic rearrangement and simplification, as well as applying the definition of capacitance.

Step	Reason
5. $\kappa = \dfrac{q}{\Delta V_\kappa}\dfrac{d}{\varepsilon_0 A}$	divide the fractions
6. $\kappa = C_\kappa \dfrac{d}{\varepsilon_0 A}$	definition of capacitance
7. $C_\kappa = \kappa\left(\dfrac{\varepsilon_0 A}{d}\right)$	solve for C_κ

28.18 - Sample problem: how big should a computer key be?



In a capacitor keyboard, each key rests on a parallel-plate capacitor filled with a compressible insulator having a dielectric constant of 5.15.

The capacitance of each key is $3.13\text{x}10^{-12}$ F when resting, and $7.86\text{x}10^{-10}$ F when pressed down.

When you type on a keyboard, you push each key down 1.73 mm. What is the area of the capacitor plate?

To solve this problem you will use your knowledge of how the capacitance of a parallel-plate capacitor depends on its dimensions and on the dielectric constant of the insulator that separates the plates.

Variables

resting capacitance	$C_\kappa = 3.13\times10^{-12}$ F
capacitance when key depressed	$C_\kappa{}' = 7.86\times10^{-10}$ F
area of each capacitor plate	A
resting plate separation	d
change in plate separation	$\Delta d = 1.73\times10^{-3}$ m
dielectric constant of insulator	$\kappa = 5.15$

Strategy

1. Write two equations for capacitance: one for the resting position and one for the depressed position.
2. Solve the pair of equations for the area of the capacitor plates. You do not need to find the value of the other unknown, the resting plate separation.

Physics principles and equations

The capacitance of a parallel-plate capacitor containing a dielectric with constant κ is

$$C_\kappa = \kappa\left(\frac{\varepsilon_0 A}{d}\right)$$

Step-by-step solution

First we write the equations for the resting and depressed capacitances.

Step	Reason
1. $C_\kappa = \kappa\left(\frac{\varepsilon_0 A}{d}\right)$	resting capacitance
2. $d = \frac{\kappa\varepsilon_0 A}{C_\kappa}$	solve for d
3. $C_\kappa' = \kappa\left(\frac{\varepsilon_0 A}{d - \Delta d}\right)$	depressed capacitance

Now we solve the two capacitance equations together, solve for the area, and evaluate.

Step	Reason
4. $C_\kappa' = \dfrac{\kappa\varepsilon_0 A}{\left(\frac{\kappa\varepsilon_0 A}{C_\kappa}\right) - \Delta d}$	substitute equation 2 into equation 3
5. $A = \dfrac{C_\kappa C_\kappa' \Delta d}{\kappa\varepsilon_0 (C_\kappa' - C_\kappa)}$	solve for A
6. $C_\kappa' - C_\kappa = (7.86\times10^{-9}\text{ F}) - (3.13\times10^{-11}\text{ F})$ $C_\kappa' - C_\kappa = 7.83\times10^{-9}$ F	evaluate $C_\kappa' - C_\kappa$
7. $A = \dfrac{(3.13\times10^{-12}\text{F})(7.86\times10^{-10}\text{F})(1.73\times10^{-3}\text{m})}{(5.15)(8.85\times10^{-12}\frac{\text{C}^2}{\text{N}\cdot\text{m}^2})(7.83\times10^{-10}\text{F})}$ $A = 1.19\times10^{-4}\text{ m}^2$ $A = 1.19\text{ cm}^2$	evaluate area

This is a little over one square centimeter, slightly less than the size of a keyboard key.

28.19 - Gotchas

The total amount of charge on the plates of a capacitor is designated by q. No, the amount of charge on **one** plate of the capacitor is q. Each plate has the same amount of charge: positive on one plate and negative on the other. The total charge is zero.

Capacitors have a net charge. No, the two plates have equal and opposite amounts of charge, so the capacitor is electrically neutral. When you say "the capacitor has a charge" or the like, it means each plate is charged.

The energy of a charged capacitor can be computed by calculating how much work a source of potential difference must do on each electron that moves from the positive to the negative capacitor plate. This idea only works as a thought experiment, which is actually based on infinitesimal increments of negative charge rather than electrons. In reality, a capacitor gets charged as electrons get moved through a circuit that connects its plates. Please do not think charge flows between the plates because of this thought experiment!

28.20 - Summary

A capacitor is a device with two conductors that are placed in close proximity, and which store equal but opposite amounts of charge.

In an ideal parallel-plate capacitor, the conductors consist of two flat conducting plates separated by a small distance.

Other kinds of capacitors are cylindrical capacitors and spherical capacitors.

Capacitors are characterized by their capacitance, represented by the symbol *C*, which equals the amount of charge on one conductor divided by the potential difference between the conductors. The unit of capacitance is the farad (F).
1 F = 1 C/V.

By storing charge, capacitors also store electric potential energy. A battery or other device causes charge to accumulate on the conductors of the capacitor. The opposite charges on the conductors create an electric field between them. Electric potential energy is stored in the electric field.

Electric field energy density is the amount of electric potential energy stored in an electric field, per unit volume. It is proportional to the square of the electric field strength.

Between the conductors of a capacitor there may be a dielectric, an insulator that can increase the capacitance by decreasing the field strength between the conductors for a given charge. The factor characterizing the increase in capacitance (or the decrease in field strength) is called the dielectric constant, κ.

Equations

Definition of capacitance

$$C = \frac{q}{\Delta V}$$

Parallel-plate capacitor

$$C = \frac{\varepsilon_0 A}{d}$$

Cylindrical capacitor

$$C = \frac{2\pi\varepsilon_0 h}{\ln(b/a)}$$

Spherical capacitor

$$C = \frac{4\pi\varepsilon_0 ab}{b - a}$$

Potential energy in a capacitor

$$PE_e = \frac{q^2}{2C}$$

$$PE_e = \frac{C(\Delta V^2)}{2}$$

Electric field energy density

$$u_E = PE_e / Vol$$

$$u_E = ½\, \varepsilon_0 E^2$$

Dielectric constant

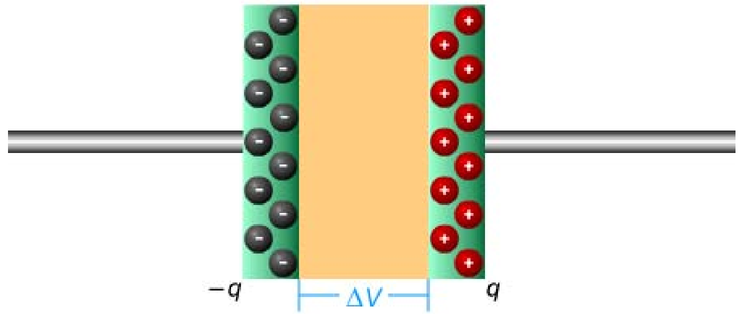
$$\kappa = E_{vacuum} / E_{dielectric}$$

$$C_\kappa = \kappa C$$

Chapter 28 Problems

Conceptual Problems

C.1 If you increase the potential difference between the plates of a capacitor, what happens to (a) the amount of charge on each plate and (b) the **net** charge of the capacitor?



(a) i. It increases
ii. It decreases
iii. It stays the same
(b) i. It increases
ii. It decreases
iii. It stays the same

C.2 A dielectric is inserted between the plates of an isolated, charged capacitor, replacing the vacuum that was there before. Its dielectric constant is greater than one. (a) Has the capacitance increased, decreased, or stayed the same? (b) Is there more, less, or the same amount of charge on the capacitor?

(a) i. Increased
ii. Decreased
iii. Stayed the same
(b) i. More charge
ii. Less charge
iii. Same amount of charge

C.3 You insert a dielectric into an isolated, charged capacitor. Does the potential energy stored by the capacitor increase, stay the same, or decrease during the insertion?

i. It increases.
ii. It stays the same.
iii. It decreases.

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to determine whether the charge on a capacitor is proportional to the potential difference.

i. Charge is proportional to potential difference
ii. Charge is not proportional to potential difference

Section 1 - Capacitors

1.1 Suppose you disassemble a cell phone and remove a capacitor inside labeled " 6.5e−11 F." If you apply a potential difference of 5.0 volts across the capacitor, how much charge would you expect to be on the positive plate?

__________ C

1.2 When a 9.0 V battery is connected to a capacitor, it puts 6.0e−4 C of charge on the capacitor. When a different battery with an unknown voltage is connected to the same capacitor, it puts 1.5×10^{-4} C of charge on the capacitor. What is the unknown battery's voltage?

__________ V

1.3 A capacitor accumulates 6.4e−5 C of charge when connected to a 9.0 volt battery. What is its capacitance?

__________ F

1.4 A 6.1e−6 F capacitor accumulates 51 μC of charge when connected to a battery. What is the battery's voltage?

__________ V

1.5 When a battery is connected to a 2.0 microfarad capacitor, it places an equal but opposite charge on each plate, of magnitude 4.0×10^{-5} coulombs. How much would the same battery charge a 3.6e−6 farad capacitor?

__________ C

Section 3 - Gauss' law and the capacitance of a parallel-plate capacitor

3.1 Two circular metal plates with area 0.28 square meters are placed 1.6e−3 m apart to form a parallel-plate capacitor. Find its capacitance.

__________ F

3.2 The plates of a 5.50e−9 F parallel-plate capacitor are separated by 1.20 mm. What is the area of each plate?

__________ m^2

3.3 A 665 pF parallel-plate capacitor has rectangular plates that measure 0.550 m by 0.440 m. Find the distance between the plates.

__________ m

3.4 The plates of a capacitor are separated by 0.0060 m. The capacitor is charged by a potential difference of 5.0 V and then disconnected. Now you separate the plates to a distance of 0.015 m. What is the new potential difference between the plates?

__________ V

Section 5 - Interactive problem: capacitor plate separation

5.1 Use the information given in the interactive problem in this section to answer the following question. What must the separation of the plates be to achieve a 2.00 nC charge on a plate? Test your answer using the simulation.

__________ m

Section 7 - Energy in capacitors

7.1 In an electronic circuit, you need a capacitor to store 4.50e−9 J of energy. You have 1.50 volts available to charge it with. What capacitance should you choose?

__________ F

7.2 Suppose a capacitor in a video camera has a capacitance of 26.0 μF, and is currently holding 5.20e−5 C of charge. How much potential energy is stored in the capacitor?

__________ J

7.3 How much potential energy is stored in a 2.10e−6 F capacitor charged by a potential difference of 651 V?

__________ J

7.4 A warning label on a 56.1 μF capacitor says that the capacitor can safely hold no more than 5.00 J of energy. What is the maximum potential difference you can safely apply to the capacitor?

__________ V

7.5 A capacitor that initially holds a charge of 3.31×10^{-3} C on its plates fully discharges, in the process doing 5.50 J of work. What is its capacitance?

__________ F

7.6 A 79.1 μF capacitor is initially uncharged. After you expend 263 J of energy to charge it up, how much charge does it hold on each plate?

__________ C

7.7 A 6.40 nF capacitor is initially charged by a potential difference of 1.40×10^3 V. Later, the potential difference is increased so that 5.10 μC of charge are added. What is the final electric potential energy stored in the capacitor?

__________ J

Section 9 - Interactive problem: You are Zeus

9.1 Use the information given in the interactive problem in this section to calculate the potential difference that will give the lightning bolt the correct amount of energy to light the Olympic flame, and start the games. Test your answer using the simulation.

__________ V

Section 10 - Capacitors and electric field energy density

10.1 0.0031 J of energy are stored in a parallel-plate capacitor whose plates each have area 0.30 m^2 and are 1.5 mm apart. What is the electric field energy density between the plates?

__________ J/m^3

10.2 A parallel-plate capacitor whose plates are separated by 2.00 mm is charged by a 455 volt potential difference. What is the energy density of the field between the capacitor plates?

__________ J/m^3

10.3 A parallel-plate capacitor consists of two large circular sheets of metal separated by 1.53 cm. The electric field energy density between plates is 4.41 J/m^3 while the capacitor holds 0.127 J of potential energy. What is the radius of the circular plates?

__________ m

10.4 You measure the electric field energy density inside a parallel-plate capacitor to be 0.67 J/m^3. What is the strength of the electric field?

__________ N/C

10.5 A parallel-plate capacitor is charged with 0.220 C. If the area of each of its plates is 0.620 m^2, what is the electric field energy density between the plates?

__________ J/m^3

Section 11 - Physics in medicine: defibrillator

11.1 A certain defibrillator sends 0.12 C of charge through the heart. The capacitor responsible for holding this charge before it is released has a capacitance of 95 μF. How much energy is released when it discharges?

__________ J

Section 13 - Dielectrics

13.1 The dielectric paste used in a tiny 360 nF capacitor has a dielectric constant of 50. What would the capacitor's capacitance be without the dielectric?

__________ nF

13.2 An isolated parallel-plate capacitor with no dielectric has an electric field of 8.4×10^3 V/m. When you inject dielectric paste from an unlabeled jar, the field goes down to 1.4×10^3 V/m (the charges on the plates of the capacitor stay the same). What is the dielectric constant of the unlabeled paste?

____

13.3 A tiny capacitor found in a clock radio uses a dielectric with a dielectric constant of 262. The electric field strength between the "plates" of this capacitor is 6.11×10^5 V/m. If the dielectric were replaced with a vacuum, what would the strength of this electric field become, assuming the charges on the plates remain the same?

__________ V/m

13.4 When you insert a sheet of wood (κ = 2.00) between the plates of a capacitor, the capacitance becomes 22.0 nF. What capacitance will it have if you insert a block of glass (κ = 6.00) instead?

__________ nF

Section 16 - Physics at work: commercial capacitors

16.1 A variable capacitor has a range of capacitances from 2.0 nF to 20 nF. What is the difference between the most and least charge that can be stored on the variable capacitor when it is connected to a 9.0 V source of potential difference?

__________ C

Section 17 - Derivation: capacitance with a dielectric present

17.1 A parallel-plate capacitor consists of plates of area 0.44 m^2 separated by 1.0 mm of vacuum. The capacitor is submerged in water (κ = 80.2) and water fills the space between the plates. What is its capacitance now?

__________ F

Additional Problems

A.1 You create a capacitor by sandwiching 33 sheets of 8.50 by 11.0 inch paper between two pieces of similarly-sized aluminum foil. This paper has a dielectric constant of 3.65, and a stack of 500 sheets is 4.97 cm tall. What is the capacitance of your contraption?

__________ F

A.2 You want to store enough energy in a 1.60 F capacitor such that if the energy were released at a uniform rate, it could run a 60.0-watt lightbulb for 1.00 minutes. What potential difference should you use to create the initial charge on the capacitor?

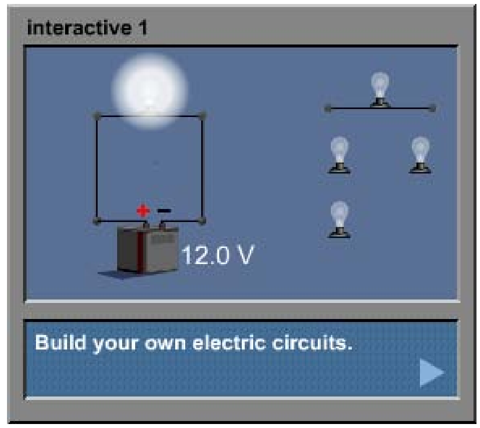
__________ V

A.3 You have two parallel-plate capacitors, one with square plates and the other with circular plates. In both capacitors the distance between plates is 1.50 mm. The circular plates have a diameter of 0.500 m, and the square plates have a width and length of 0.500 m. By inserting dielectrics, you want to make the capacitance of the two capacitors equal. If the dielectric constant of the material in the square capacitor is 6.40, what must the dielectric constant of the material in the circular capacitor be?

____

chapter 29 Direct Current Circuits

29.0 - Introduction



Electric circuits include components such as batteries, resistors and capacitors. These basic elements can be combined in a myriad of ways. How these components function, by themselves and in combination, defines the fundamental operation of electric circuits. In this chapter, we examine direct current electric circuits, circuits in which the current always flows in the same direction.

In this chapter, you will learn how to analyze direct current circuits. On the right is a simulation in which you can build your own electric circuits. Initially, the circuit consists of a battery, a light bulb and wires that connect these components. The simulation also contains additional wire segments and light bulbs.

You can place light bulbs in various places in the circuit. Pay special attention to the brightness of the light bulbs: The brighter the light, the more power the circuit is supplying to it. You can also use two devices in the control panel to study the circuit. One, a voltmeter, measures the potential difference across components, while the other, an ammeter, measures the current flowing through the wire at any location in the circuit.

In this simulation, the battery and wires effectively have zero resistance. The simulation includes a total of five light bulbs, each with a resistance of 50 ohms. You add and remove light bulbs and wires by dragging them; the components will snap into place. Only one light bulb can be placed on each wire segment.

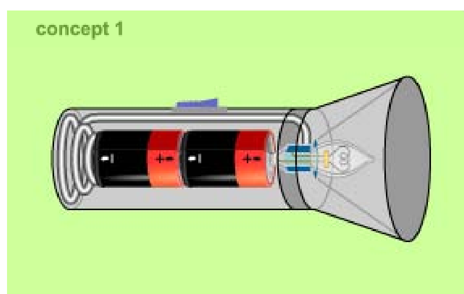
The purpose of this simulation is for you to experiment with the electric components in a circuit. One way to start is by assessing the circuit in its initial state using the voltmeter and ammeter. How does the potential difference across the battery compare with the potential difference across the light bulb? What about the current? Is it the same everywhere or does it differ from place to place?

Now add another light bulb above the first one in the circuit: Snap in two wire segments in a vertical orientation, and then put a segment containing a light bulb between them. How does the potential difference across the battery now compare to the potential difference across each light bulb? Is the current still the same everywhere? This time you should find that the current can differ by location.

You can also use the voltmeter to confirm Ohm's law. Since you are told the bulb's resistance (50 ohms) and can measure the potential difference across it using the voltmeter, you can use the law to calculate the current flowing through the wire segment containing the bulb. You then can verify your calculation using the ammeter.

You are probably thinking that this introduction has asked you to answer a lot of questions! If you cannot answer all the questions above, that is fine. This chapter is dedicated to preparing you to address them.

29.1 - Electric circuits



Electric circuit
Set of electric components connected by wires

In this section, we provide an overview of the components of a basic circuit. All these components merit more discussion, but here we want to provide some context on how they function together in a circuit.

A flashlight provides an example of a circuit. The flashlight we show contains two batteries, a switch, a light bulb, and some metal wires that connect these components.

In the flashlight, the batteries are the source of the energy that causes the net motion of charge in the circuit. The ends of a battery are at different electric potentials. There is a potential difference across the two terminals that are at opposite ends of the battery. A typical potential difference for a battery like those shown is 1.5 volts. The terminal with the greater electric potential is marked with a plus (+) sign, and the terminal with the lower electric potential is marked with a negative (−) sign. Putting together two batteries as shown creates a potential difference across the two batteries of approximately 3.0 volts. The batteries are used to create an electric field that causes a net flow of electrons: a current.

Current will only flow when there is a complete path: a loop, as opposed to a dead end. (Circuit comes from the Latin word *circumire*, to go around). When the switch is in the "off" position, it creates a gap in the circuit, and current cannot flow. When the switch is pushed to "on", the gap is closed and a current can flow.

We have highlighted the circuit inside a flashlight in Concept 2. Let's trace the direction of conventional current around the circuit. A positive charge starts at the positive terminal of the battery on the right, and moves through a coiled wire called the filament in the light bulb. It exits the filament and moves through the wire that contains the switch. The charge then flows through the batteries and starts its round trip over again.

There is resistance inside the batteries, in the wires and in the light bulb. The resistance inside the batteries and in the wires is minor compared to that of the filament in the light bulb, and it is often reasonable to ignore these minor resistances. The light bulb is the major source of resistance, and it supplies what is called the *load resistance* of the circuit. The resistance of this component and the potential difference across the batteries determine the amount of current in the circuit.

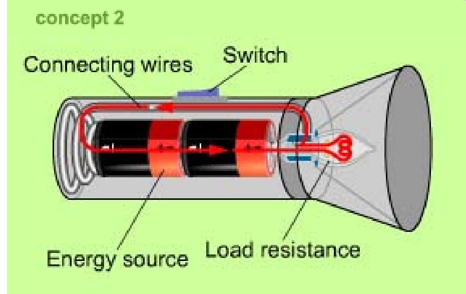
The flashlight creates light when current flows through the filament. Modern day filaments often are made of very thin tungsten wires. As a current passes through the filament, electrons lose energy in collisions with the atoms of the filament, and the filament's temperature increases. It becomes hot (up to $3000°C$), and energy leaves the filament in the forms of heat and light.

Although tungsten is an effective filament, when it is hot the tungsten vaporizes and the filament becomes thinner. This increases its resistance, which means it shines less brightly. Eventually, when it "burns out", it becomes so thin and brittle that it ruptures.

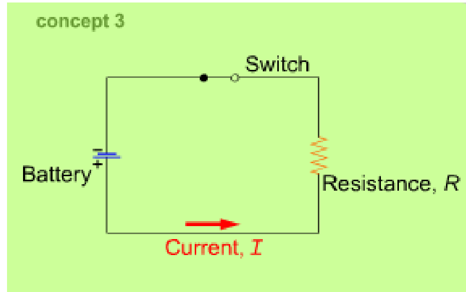
The flashlight contains all the essential elements of a circuit. It has a source of energy, the batteries. They create a potential difference that causes electrons to move. As the electrons move through the light bulb, they encounter resistance, which causes this resistor to dissipate energy.

The current flows only in one direction in this circuit, which makes it a *direct current* circuit. Direct current circuits are the topic of this chapter. We study these circuits primarily when the current has reached a steady state − a constant flow − not in the brief moments when the current changes, such as immediately after the flashlight is switched on or off.

Scientists and engineers use symbols to represent components in circuit diagrams, as shown in Concept 3. The thin black lines represent wires. The battery and resistance symbols are labeled in the circuit. The switch is "on", so it is in its closed position in this diagram. The red arrow indicates the direction of flow of the conventional current.



Circuits usually contain:
An energy source
A load resistance
Wires connecting it all in a loop
A switch



Drawing electric circuits
Use circuit diagrams with symbols

29.2 - Electromotive force

Electromotive force (emf, $\mathcal{E}$): Maximum potential difference from an energy source such as a battery.

Batteries are a common source of emf.

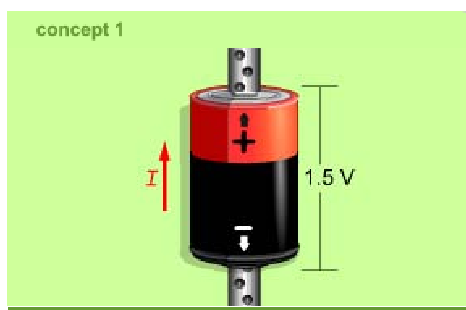
A potential difference applied across a conductor causes a current to flow. Common devices such as a flashlight need a continuing source of potential difference − an *emf* − so that current will keep flowing and the flashlight's bulb will stay illuminated.

A battery often supplies the emf. The symbol for emf is $\mathcal{E}$. Like any potential difference, an emf is measured in volts. There are many sources of emf: electric generators, solar photovoltaic cells and so on. Even living creatures can be an emf source. Humans rely on emfs generated in the body to cause electric currents in nerves.

The term "electromotive force" is misleading because an emf is **not** a force. It is a potential difference and its unit is the volt. It is a well-established term in physics, however, and we will use it too. In any case, since we typically write it in its abbreviated form as emf, you should not too often be confused by seeing the word "force".

A battery is the typical emf source for direct current circuits. Chemical reactions within the battery cause one terminal of the battery to be positively charged, and the other to be negatively charged. These terminals are marked with plus (+) and minus (−) signs.

Batteries are classified by their emf. A typical battery used in a flashlight has an emf of 1.5 volts, while a car battery has an emf of 12 volts. If a battery has an emf of 1.5 volts, this means that the electric potential of its positive terminal is 1.5 volts higher than that of its negative terminal.



emf ($\mathcal{E}$)
A potential difference

Sometimes batteries are referred to as "charge pumps." They increase the potential energy of the charge flowing through the circuit. The unit of emf, the volt, equals joules per coulomb. You can think of a nine-volt battery as doing nine joules of work on each coulomb of charge that flows through it.

Batteries also are defined by how much energy they can supply over their lifetimes. AAA and D batteries are both 1.5-volt batteries, but the larger D battery supplies more energy over its lifetime. This total available energy is measured in watt-hours. A small battery for watches will have about 0.1 watt-hours of total energy, while a car battery has a total available energy of about 500 watt-hours.

When current flows through a battery, it encounters resistance. This is called a battery's *internal resistance*. This means that when placed in a circuit, a battery's emf and the potential difference across its terminals are not the same. The emf is greater than the potential difference (unless the battery is being charged, in which case the emf can be greater). The internal battery resistance is usually minor, however. Many times in this chapter we will treat it as zero, and consider the emf and the potential difference to be the same.

29.3 - Energy and electric potential in a circuit

We will use the simple circuit shown to the right as our starting point for discussing how to analyze circuits. Specifically, it is important to correctly determine the changes in electric potential across various components in the circuit.

The circuit in Concept 1 contains a resistance-free battery and a resistor: a flashlight bulb. As the diagram reflects, the resistor has a resistance of 5.0 ohms, and the battery has an emf of 1.5 volts. We want to determine the potential difference across the resistor.

To analyze the changes in electric potential occurring around the circuit, consider some charge conducting a hypothetical journey around the circuit loop, as shown in Concept 2. Imagine 0.5 coulombs of positive charge traveling one complete loop in the direction of conventional current (clockwise in this case). What happens to the potential energy of this charge as it passes through each component − the battery and the resistor − in completing a round trip around the loop?

For starters, the potential energy must be the same at the beginning and end of the closed path, because the electrostatic force is a conservative force. Just as your gravitational PE is the same after you take a round trip walk, even if you go uphill and down on the way, so too the charge's electrostatic PE is the same when it returns to its starting point.

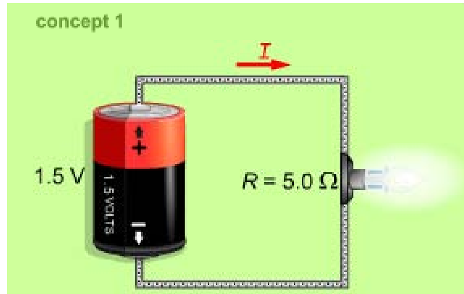
To determine the changes in potential energy as the charge makes its journey through each component, we will use an equation that relates the change in potential energy to charge and potential difference: $\Delta PE = q\Delta V$.

We will start with the charge at the negative terminal of the battery, and have it flow through the battery to the positive terminal. The **change** in potential energy for a positive charge traveling across the battery in this direction is positive. The battery must do work on this positive charge to move it toward the positive terminal. The charge's change in PE is the product of the charge and the potential difference, (0.5 C)(1.5 V), which equals +0.75 J. The electric potential energy of the charge increases and the chemical potential energy of the battery decreases correspondingly.

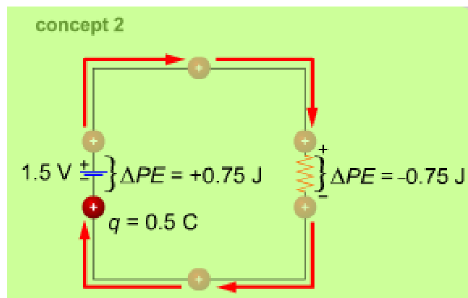
Next, the charge flows through the wire around the circuit. We assume that the wire has negligible resistance. To complete a circuit, the charge must pass through the resistor. This is the one location in this simple circuit where the charge loses energy before returning to its starting point. The resistor heats up, giving off this energy as heat. The charge's change in potential energy in the resistor must equal −0.75 joules, because the sum of the changes in potential energy around the complete loop must equal zero.

Now that we have assessed the potential energy changes around a circuit loop, we can also assess the changes in electric potential. Concept 3 shows the changes in electric potential around a complete circuit loop. The changes in electric potential can be determined using the equation $\Delta V = \Delta PE/q$.

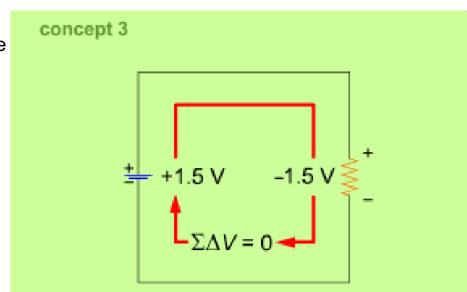
Concept 3 again shows a clockwise path around the circuit. Moving in this direction, the electric potential increases by +1.5 V as the battery is traversed. The potential difference across the battery traveling in this direction is equal to the change in PE, +0.75 J, divided by the amount of charge, 0.5 C. Across the resistor, since the change in PE is −0.75 J in this direction, the change in potential is −1.5 V. The electric potential is greater on the "upstream" end of the resistor than on the "downstream" end. The charge moves across the resistor from a region of a higher electric potential to one of lower electric potential.



Battery-resistor circuit



Circuit energy analysis
PE of charge is unchanged after any round-trip
· Energy increases in battery
· Energy decreases in resistor



Electric potential changes in circuit
Potential changes around circuit sum to zero

What we have shown is fundamental and important. In this two-component circuit, the change in electric potential across one component − the battery − is equal to but opposite the change in electric potential across the other component: the resistor. The changes sum to zero. We have shown this in a specific case, but the rule holds in general for any closed loop around a circuit.

In fact, you are learning an important principle about electric potential changes in any circuit loop. It is called Kirchhoff's loop rule, and is discussed in more depth in another section.

The direction we chose to travel the circuit was arbitrary. Had we chosen to travel the circuit in a counterclockwise direction, the change in electric potential would be negative as we crossed the battery and positive as we crossed the resistor. We also chose to consider a positive charge rather than a negative one. In any case, clockwise or counterclockwise, positive or negative, the principle holds that the changes in electric potential around a complete circuit loop sum to zero.

This section can be used to illustrate a more general principle: As you traverse a circuit, you can add the potential differences across the components. For instance, if you traversed two resistors, the potential difference across the two resistors would equal the sum of each resistor's potential difference.

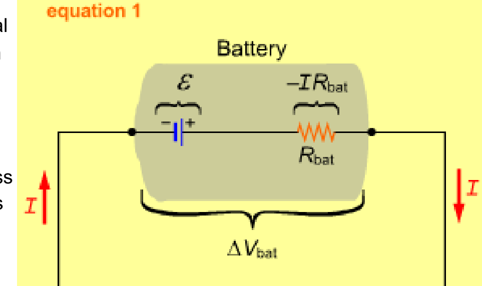
29.4 - Internal resistance of a battery

Batteries have internal resistance. It is often convenient and a good approximation to ignore internal resistance, as we do through much of this chapter. However, the internal resistance in batteries has a small but measurable effect on circuits, so we discuss it in more detail here.

To calculate the potential difference across the terminals of a battery, we sum the battery's emf $\mathcal{E}$ and the potential difference ΔV_R caused by the internal resistance of the battery. In the direction of conventional current, the change in potential ΔV_{bat} across the battery is positive, while the change in potential caused by the internal resistance is negative. Equation 1 shows this in equation form.

When no current flows, the battery's internal resistance is zero and the emf equals the potential difference across the battery's terminals. As the current increases, the internal resistance increases, the potential difference caused by this internal resistance increases, and the potential difference across the battery decreases.

equation 1

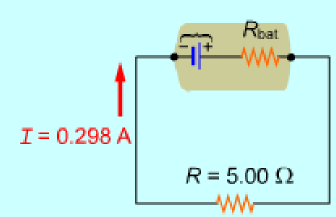


Potential difference across battery

$$\Delta V_{bat} = \mathcal{E} - \Delta V_R = \mathcal{E} - IR_{bat}$$

ΔV_{bat} = potential difference across battery
$\mathcal{E}$ = emf, I = current
ΔV_R = resistive potential difference
R_{bat} = internal resistance of battery

example 1



The battery's emf is 1.50 V and its internal resistance is 0.0336 Ω. What is the potential difference across its terminals?

$\Delta V_{bat} = \mathcal{E} - IR_{bat}$
ΔV_{bat} = 1.50 V – (0.298 A)(0.0336 Ω)
ΔV_{bat} = 1.49 V

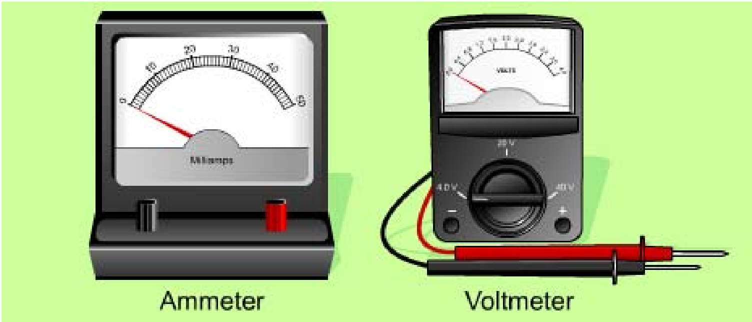
29.5 - Measuring current and potential difference

Ammeter: A device that measures current in a wire.

Voltmeter: A device that measures the potential difference between two points.

Ammeters and voltmeters are common tools for analyzing circuits. Both are shown above.

Ammeters measure current. An ammeter is inserted in the circuit, as shown in Concept 2, so that all the current in that part of the circuit flows through it. Because current flows through them, ammeters are built to have extremely low resistance in order to minimize their impact on the circuit. An ideal ammeter would have no resistance.



concept 1

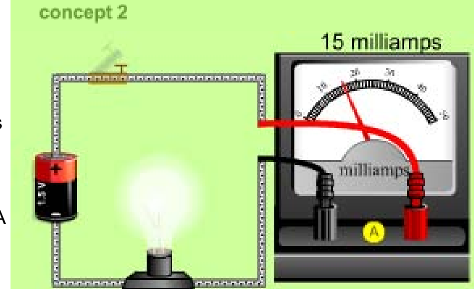
Ammeters and voltmeters

A voltmeter measures the potential difference (voltage) between two points in a circuit. In Concept 3, the potential difference across the two terminals of a battery is being measured. To use the voltmeter, a lead is placed on each side of the battery.

Voltmeters are also designed to have minimal effects on a circuit. In the case of the voltmeter, this means they are designed to have a high resistance, so little current flows through them, and the rest continues to flow through the circuit component being measured.

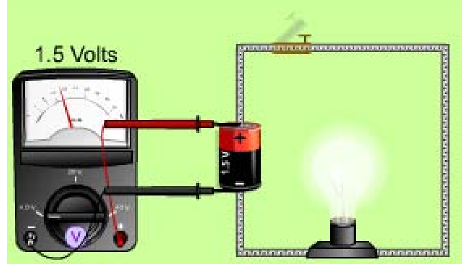
Ammeters and voltmeters often are combined into one instrument called a *multimeter*. A multimeter may also contain a device for measuring resistance, called an *ohmmeter*.

concept 2



Ammeter
Measures current

concept 3



Voltmeter
Measures potential difference

29.6 - Series wiring

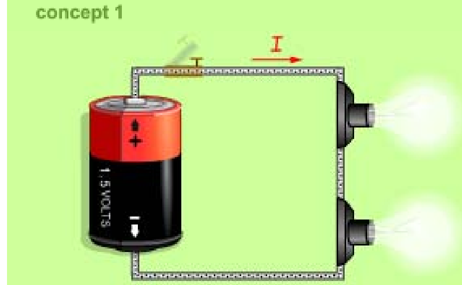
Series wiring: Circuit wiring in which the components are placed one after another. All of the current flows through each component.

The diagram to the right shows an example of a series circuit. The same amount of current passes through the battery and through each light bulb. The circuit has no branches; there are no places where the current can split to follow another path. The light bulbs and the battery are said to be connected in series.

Place an ammeter anywhere in this circuit and you will measure the same value for the current. The same amount of current passes through the battery, the first light bulb, the second light bulb and the wires that connect these components.

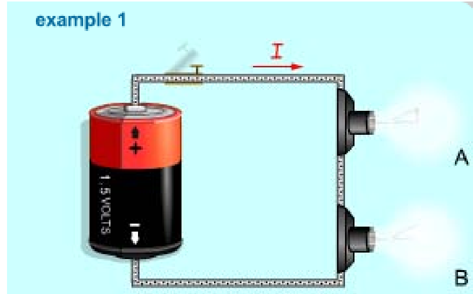
If the current in a series circuit is interrupted anywhere, it is interrupted everywhere. Because there are no alternative routes, if the circuit is broken at any point, then there is no detour path for the current to follow.

concept 1



Series wiring
Current has one path through components
Current same at all points

This phenomenon is explored in more detail in Example 1. Two light bulbs are wired in series, similar to old-fashioned holiday lights. The filament of each light bulb is part of the circuit. When a light bulb burns out, its filament breaks, and there is no path for current through that bulb. The example problem asks what happens to the second light bulb in a series circuit if the first bulb burns out.



If light bulb A burns out, will light bulb B remain lit?

No – a break in a series circuit causes all current to stop

29.7 - Resistors in series

To calculate the equivalent resistance of resistors in series: Add each resistor's resistance.

Analyzing circuits with multiple components can be complex. The task can sometimes be simplified by treating several components of the same type as if they were one. You can simplify a circuit by calculating what is called the "equivalent resistance" or "equivalent capacitance" of multiple resistors or capacitors. You can then treat the components as one, using their equivalent resistance or equivalent capacitance.

Our first case for doing this is resistors in series. Once we determine their equivalent resistance, we can treat them as though they were one component, and determine how much current flows through the circuit in Example 1. In the case of resistors wired in series, the equivalent resistance is the sum of the resistances. We show this in Equation 1.

Consider the circuit shown in Example 1 that contains two resistors in series. The question asks for the amount of current in this circuit. You can determine the current using Ohm's law, $\Delta V = IR$, but what should you use for R?

First, determine the series circuit's equivalent resistance by summing the individual resistances. The 6 Ω resistor plus the 4 Ω resistor equals an equivalent resistance of 10 Ω. The circuit in the example has a 20-volt battery. Since the two resistors create an equivalent resistance of 10 Ω, you can use Ohm's law to calculate the current through them. It equals 20 V divided by 10 Ω, or 2 A.

Derivation. Why can resistances in series be added? We now derive the series rule for resistors, for n resistors in series.

Variables

potential difference across all resistors	ΔV_{total}
potential difference across i^{th} resistor	ΔV_{i}
current through circuit	I
current through i^{th} resistor	I_{i}
equivalent resistance of circuit	R_{equiv}
resistance of i^{th} resistor	R_{i}

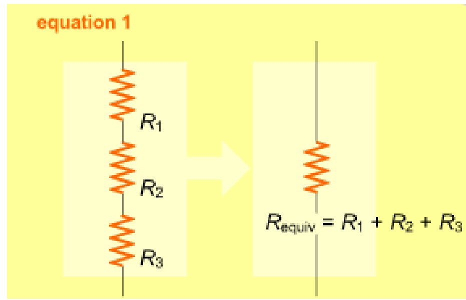
Strategy

1. Find the total potential difference across all the resistors combined.
2. Use Ohm's law to rewrite potential difference in terms of current and resistance. An algebraic simplification gives the series rule for resistors.

Physics principles and equations

With components in series, the total potential difference equals the sum of the potential differences across each component.

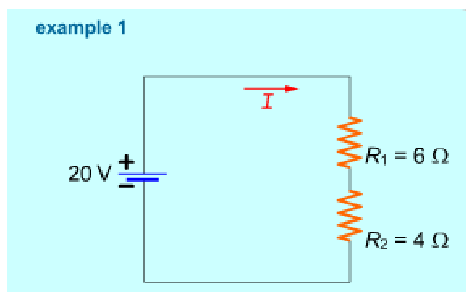
Ohm's law relates potential difference, current and resistance:



Resistors in series

$$R_{\text{equiv}} = R_1 + R_2 + \ldots + R_{\text{n}}$$

Equivalent resistance = sum of individual resistances



What is the equivalent resistance of the resistors and what is the current?

$R_{\text{equiv}} = R_1 + R_2$

$R_{\text{equiv}} = 6\ \Omega + 4\ \Omega = 10\ \Omega$

$$I = \frac{\Delta V}{R}$$

$$I = \frac{20\ \text{V}}{10\ \Omega} = 2\ \text{A}$$

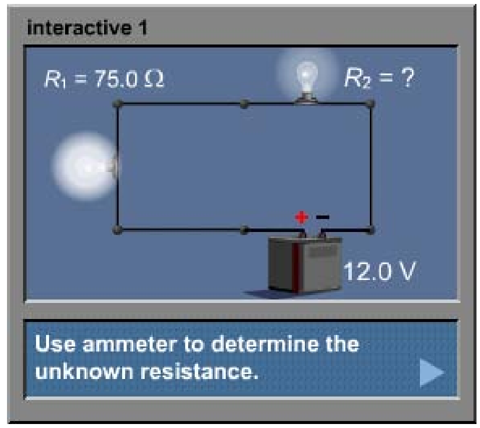
$$\Delta V = IR$$

With components in series, the same amount of current flows through each component.

Step-by-step derivation

Step	Reason
1. $\Delta V_{\text{total}} = \Delta V_1 + \ldots + \Delta V_n$	total ΔV is sum of series ΔV's
2. $IR_{\text{equiv}} = I_1R_1 + \ldots + I_nR_n$	Ohm's law
3. $I = I_1 = I_2 = \ldots = I_n$	current the same
4. $R_{\text{equiv}} = R_1 + \ldots + R_n$	divide by equal current

29.8 - Interactive problem: series wiring



The circuit to the right contains two light bulbs and a 12-volt battery. One of the light bulbs, R_1, has a resistance of 75 Ω. Your task is to determine the resistance of the R_2 bulb.

You have an ammeter at your disposal. You drag this tool onto the circuit to measure the current.

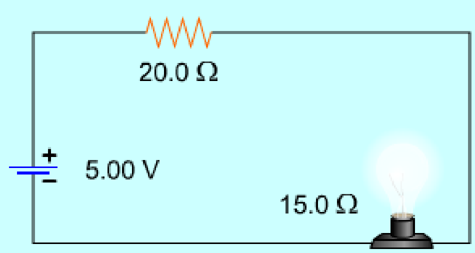
Use the ammeter, Ohm's law and your knowledge of equivalent resistance in series circuits to determine the resistance of the R_2 bulb. In order to solve the problem you may want to first consider what the equivalent resistance of the two resistors must be given the reading of the ammeter and the voltage shown for the battery.

Type your answer in the space provided. Press CHECK, and a message will indicate whether your answer is correct.

You can try again by entering a new answer and pressing CHECK again.

If you have trouble solving this problem, review Ohm's law and the section on the equivalent resistance of resistors in series.

29.9 - Interactive checkpoint: voltage divider



A *voltage divider* is often used in circuits where the potential difference available is too large for certain components. A resistor is placed in series with the sensitive component, "using up" some of the potential difference available. In the circuit shown, a 5.00 V battery is connected to a 20.0 Ω resistor and a sensitive 15.0 Ω light bulb. What is the potential difference across the light bulb? Use equivalent resistance to solve the problem.

Answer:

V = ______ V

29.10 - Parallel wiring

Parallel wiring: Circuit wiring that branches. The same potential difference exists across each branch.

In parallel wiring, there are junctions where multiple wires come together. A current can flow into a junction and then divide along different paths. A path between two junctions is called a *branch*.

Consider electrons just to the left of junction A in Concept 1. The electrons are moving to the right when they reach the junction. As they reach it, some turn left into the branch along the middle wire, while others continue moving straight into the branch around the outer wire loop. The two flows of electrons, the two currents, rejoin at the junction labeled B.

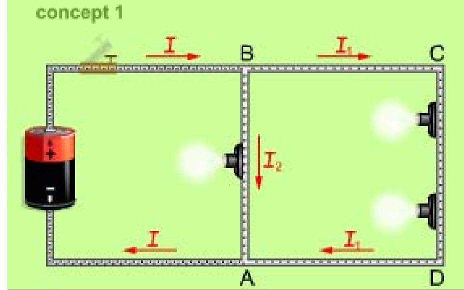
Because charge is conserved around a circuit, the sum of the currents flowing into a junction equals the sum of the currents flowing out. This is an important principle known as Kirchhoff's junction rule.

The potential difference is the same across the end points of parallel branches in a circuit. This is a crucial concept required for understanding the functioning of parallel circuits.

In the circuit on the right, a battery is connected in parallel with the light bulb in the branch AB, and with the light bulbs in the branch CD. The potential difference is identical across the battery and these two branches. The potential differences do not sum as in a series circuit.

In this circuit, the battery is a 1.5-volt flashlight battery. If you placed a voltmeter's leads on either side of the battery, you would read a value of 1.5 volts. You would also read the same value if you placed the leads across the middle branch or across the CD branch. The potential difference across all three branches is identical. This confirms that all three branches are in parallel.

concept 1



Parallel wiring
Current has more than one path
· Current may vary from loop to loop
Potential difference same across parallel branches

29.11 - Resistors in parallel

To calculate the equivalent resistance of resistors in parallel: Add the reciprocal of each resistor's resistance. The reciprocal of this sum equals the equivalent resistance.

The equation on the right shows how to calculate the equivalent resistance of resistors in parallel. First, take the reciprocal of each resistance. Those values are summed. The reciprocal of that sum is the equivalent resistance of the parallel resistors.

Another equation, also shown to the right, allows you to quickly calculate the equivalent resistance when just two resistors are wired in parallel. This equation can be derived from the first with a little algebra.

The example problem on the right shows two resistors, one of 4.0 Ω, the other of 6.0 Ω, wired in parallel. To calculate the equivalent circuit resistance of these two resistors, first invert each value and add these reciprocals. Then, invert that sum. The equivalent resistance is 2.4 Ω.

We can then calculate the current in the circuit. We must specify the current's location, because the current is not the same in all parts of the circuit. In the example problem, we specify that we are calculating it near the battery.

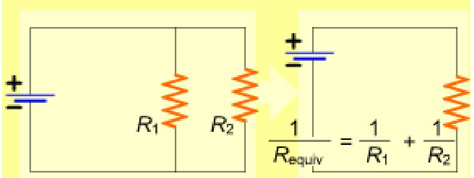
In this circuit, there is less current in the branches containing the resistors than in the branch that contains the battery. You use Ohm's law to calculate the current in the branches with the resistors. The potential difference across each branch must be the same as that of the battery, 20 volts. Using Ohm's law, you determine that there is a 5.0 A current in the branch that contains the 4.0 Ω resistor. (The current equals the potential difference, 20 V, divided by the resistance, 4.0 Ω.) A similar process enables you to determine that the current in the middle branch is 3.3 A. Since charge is conserved, you add these two currents to determine that 8.3 A flows in the branch that contains the battery.

Derivation

Variables

current through circuit	I_{total}
current through i^{th} resistor	I_{i}
potential difference across all resistors	ΔV
potential difference across i^{th} resistor	ΔV_{i}
equivalent resistance of circuit	R_{equiv}
resistance of i^{th} resistor	R_{i}

equation 1



Resistors in parallel

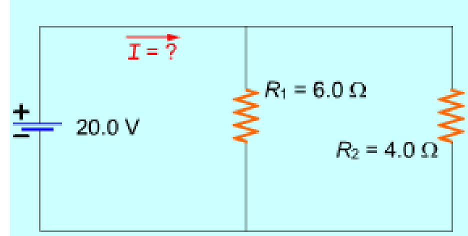
$$\frac{1}{R_{\text{equiv}}} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$$

Reciprocal of equivalent resistance = sum of reciprocals of resistances

For two resistors:

$$R_{\text{equiv}} = \frac{R_1 R_2}{R_1 + R_2}$$

example 1



What is the equivalent resistance of these resistors and what is the current at the point shown?

Strategy

1. Find the total current flowing through all the resistors combined.
2. Use Ohm's law to rewrite the current in terms of potential difference and resistance. An algebraic simplification gives the parallel rule for resistors.

Physics principles and equations

Since charge is conserved, the current flowing out of a junction equals the sum of the currents flowing into the junction.

Ohm's law

$$\Delta V = IR$$

The potential difference is the same across all parallel branches.

Step-by-step derivation

Step	Reason
1. $I_{\text{total}} = I_1 + I_2 + \ldots + I_n$	conservation of charge
2. $\dfrac{\Delta V}{R_{\text{equiv}}} = \dfrac{\Delta V_1}{R_1} + \dfrac{\Delta V_2}{R_2} + \ldots + \dfrac{\Delta V_n}{R_n}$	Ohm's law
3. $\Delta V = \Delta V_1 = \Delta V_2 = \ldots = \Delta V_n$	potential differences equal
4. $\dfrac{1}{R_{\text{equiv}}} = \dfrac{1}{R_1} + \dfrac{1}{R_2} + \ldots + \dfrac{1}{R_n}$	divide by equal potential difference

$$\frac{1}{R_{\text{equiv}}} = \frac{1}{R_1} + \frac{1}{R_2}$$

$$\frac{1}{R_{\text{equiv}}} = \frac{1}{6.0\ \Omega} + \frac{1}{4.0\ \Omega} = 0.42$$

$$R_{\text{equiv}} = \frac{1}{0.42} = 2.4\ \Omega$$

$$I = \frac{\Delta V}{R}$$

$$I = \frac{20.0\ \text{V}}{2.4\ \Omega} = 8.3\ \text{A}$$

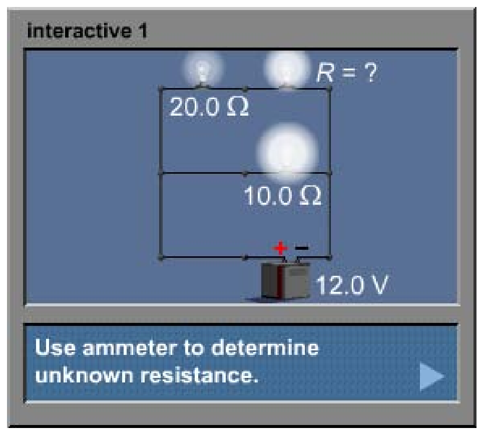
29.12 - Interactive problem: a parallel circuit

The circuit contains a battery and three light bulbs. The resistance of two of the light bulbs is known. Your task is to determine the resistance of the third light bulb.

You have an ammeter. It can be placed anywhere in the circuit to determine the current at that point. Use the ammeter, along with your knowledge of parallel circuits and Ohm's law, to determine the resistance R.

Type your answer in the space provided. Press CHECK to see if your answer is correct. You can try again by entering a new value and pressing CHECK again.

Central to solving this problem is the nature of potential difference and equivalent resistance in parallel circuits. You also need to calculate the equivalent resistance for resistors arranged in series, and apply Ohm's law. Review the sections of the textbook on these topics if you are having trouble.



29.13 - Interactive problem: potential difference in parallel circuit

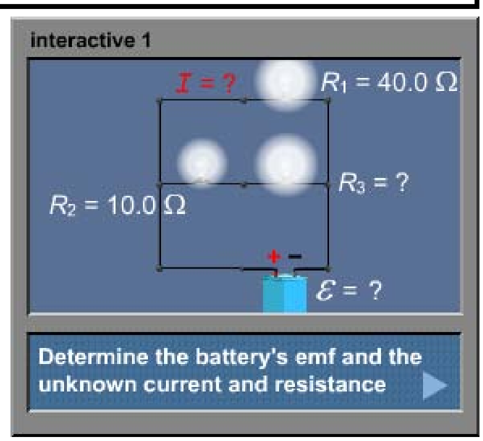
In this interactive problem, the circuit contains three light bulbs and a battery of unknown emf, as shown in the diagram to the right.

Your challenge is to use a voltmeter and your knowledge of circuits to determine the emf of the battery, the current I in the outer loop, and the resistance of the R_3 bulb.

In the simulation, the voltmeter can be placed in the circuit to determine the potential difference across two points of the circuit. However, you cannot use the voltmeter to directly determine the emf of the battery (because that would be too easy!).

Determine $\mathcal{E}$, I and R_3, and enter these values in the answer spaces provided. Press CHECK to see if your answers are correct. You can try again by entering new values and pressing CHECK again.

To solve all parts of this problem, you must understand the nature of potential difference in series and parallel circuits and how to apply Ohm's law. Review the sections of the textbook on those topics if you are having trouble.



29.14 - Circuits with series and parallel wiring

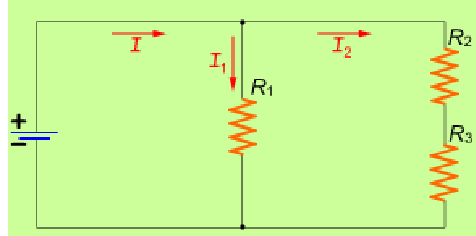
Circuits commonly include some components that are wired in series and some that are wired in parallel. The diagrams to the right show an example of such a circuit.

You can use the current to determine what is in series and what is in parallel. If the current flows along a single path from one component to another, then the two components are in series. If the current divides, then the components are in parallel.

In the circuit on the right, for example, the two resistors on the far right wire of the circuit, labeled R_2 and R_3, are in series with one another. There is no junction between them, no place for the current to split. Since there is only a single path for the current between the resistors, the resistors are in series, as emphasized in Concept 2.

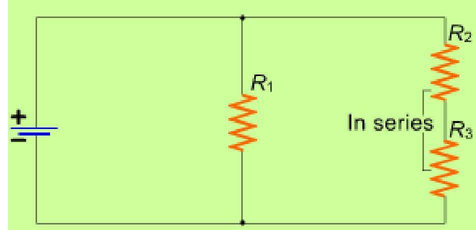
This pair of series resistors is wired in parallel with the resistor in the middle, R_1. Both branches are in parallel with the battery. Consider current that exits the battery. It encounters a junction at A where it can flow to R_1 or to the R_2R_3 combination. The current then recombines at B after flowing through those resistors. The same potential difference exists across R_1 as across the R_2R_3 combination. This means R_1 is in parallel with the R_2R_3 combination.

concept 1



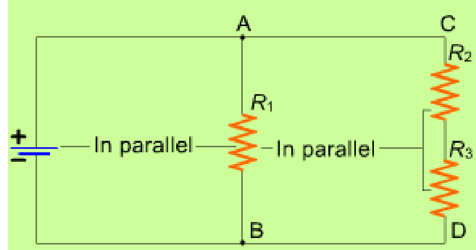
Mixed circuits
A circuit can be part series, part parallel

concept 2



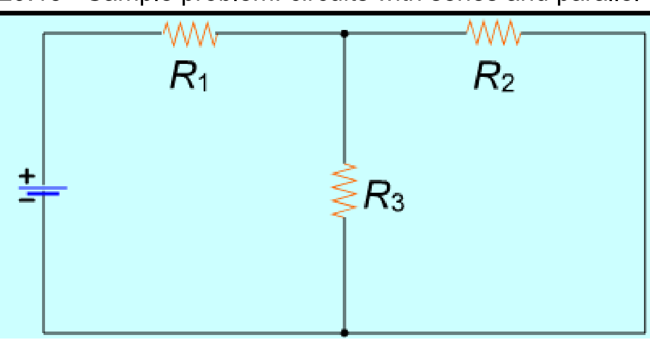
Components in series
No junction separates series components
(Single current path between them)

concept 3



Part of circuit is parallel
Junctions separate parallel components
(Current differs, potential difference the same)

29.15 - Sample problem: circuits with series and parallel wiring



Which resistors are wired in series and which are wired in parallel?

The diagram shows a circuit with three resistors and asks which resistors are wired in series and which are wired in parallel.

What is the strategy?

1. Use the definitions of series and parallel wiring.

2. Once two or more resistors are identified as series or parallel, simplify the circuit by determining the equivalent resistance. Then see whether the resulting equivalent resistance is in series or parallel with other resistors. Continue this process until the circuit cannot be simplified any further.

Physics principles and equations

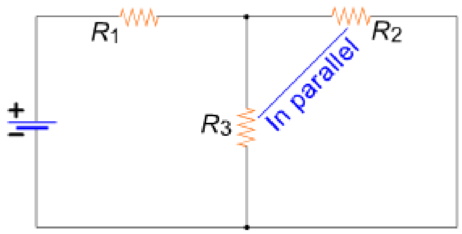
Resistors are wired in series if there is only a single path for current to flow through them.

Branches are wired in parallel if the potential difference across them is the same.

Step 1

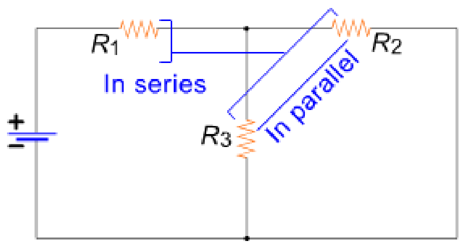
We start by checking for resistors in series. As the circuit is depicted above, none of these resistors are wired in series. The current running through any of the three resistors branches before reaching any of the other resistors.

Step 2



Next we check for parallel resistors. The resistors R_2 and R_3 are wired in parallel.

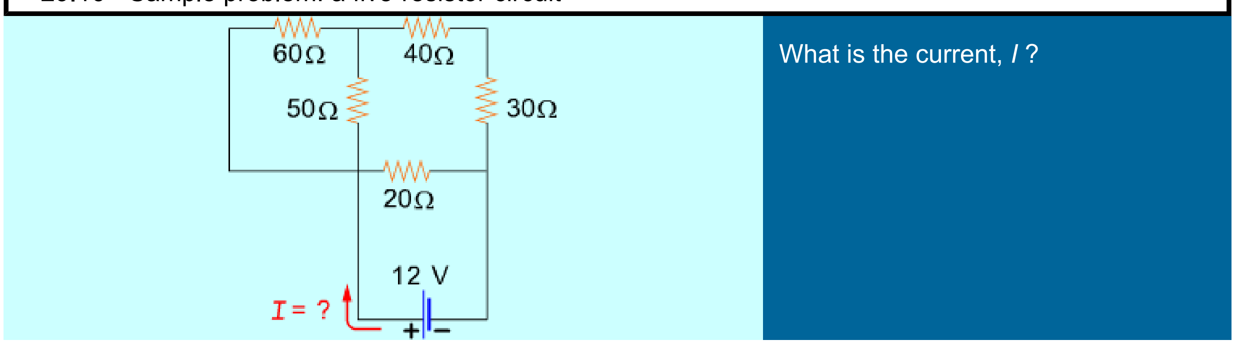
Step 3



Now, we reexamine the circuit having made a parallel combination out of R_2 and R_3. The resistor R_1 is in series with the equivalent resistance of the R_2R_3 combination. This may be tricky to see, but applying the tests for series wiring confirms it. The same current flows through R_1 as flows through the equivalent combination of R_2 and R_3. Another way of saying this is that the sum of the currents going to R_2 and R_3 is equal to the current through R_1.

(Note: It is possible to create circuits that are irreducible, in which the components are neither in parallel nor in series. If you think you have simplified a circuit as much as you can, you may be right.)

29.16 - Sample problem: a five-resistor circuit



What is the strategy?

Simplify the circuit as much as possible. Once the circuit is simplified, use Ohm's law to determine the current.

Physics principles and equations

We use the techniques for determining equivalent resistances for resistors wired in series and for resistors wired in parallel.

Ohm's law

$$\Delta V = IR$$

Equation for equivalent resistance of resistors in series

$$R_{\text{equiv}} = R_1 + R_2 + \ldots + R_n$$

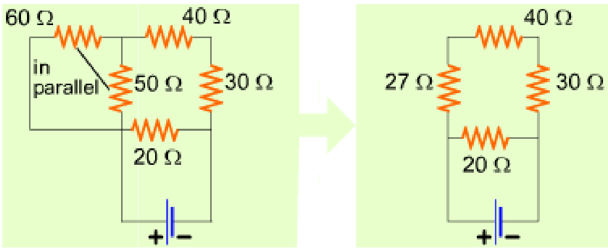
Equation for equivalent resistance of resistors in parallel

$$\frac{1}{R_{\text{equiv}}} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$$

Equation for two resistors in parallel

$$R_{\text{equiv}} = \frac{R_1 R_2}{R_1 + R_2}$$

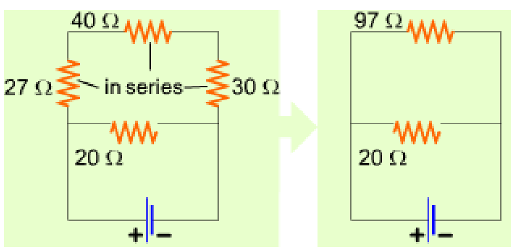
Step 1



We start by identifying two components that are in parallel: the 60 Ω and 50 Ω resistors in the upper left loop. The current branches before them and rejoins after them. These two resistors can be combined, using the steps below, resulting in the equivalent circuit shown on the right side of the diagram above.

Step	Reason
1. $R_{\text{equiv}} = \frac{R_1 R_2}{R_1 + R_2}$	equation for equivalent parallel resistors
2. $R_{\text{equiv}} = \frac{(60\ \Omega)(50\ \Omega)}{60\ \Omega + 50\ \Omega}$	enter values
3. $R_{\text{equiv}} = 27\ \Omega$	arithmetic

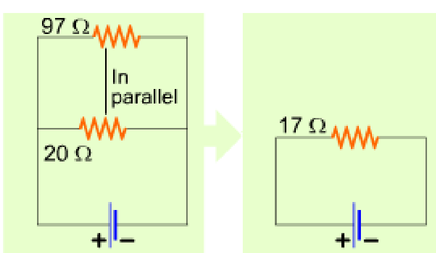
Step 2



Now we combine the resistors in series. The ones in series are shown in the diagram above. One of them is the equivalent resistance (27 Ω). They are in series because there are no junctions between them, which means the same current flows through each. (They are NOT in series with the 20 Ω resistor, which is separated from them by junctions at which the current splits.) Because they are all wired in series, we calculate the equivalent resistance by summing their resistances.

Step	Reason
1. $R_{\text{equiv}} = R_1 + R_2 + R_3$	equation for equivalent series resistors
2. $R_{\text{equiv}} = 27\ \Omega + 40\ \Omega + 30\ \Omega$	enter values
3. $R_{\text{equiv}} = 97\ \Omega$	addition

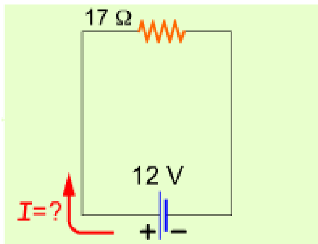
Step 3



These two resistors are wired in parallel, so we can determine their equivalent resistance by using the equation for two resistors in parallel.

Step	Reason
1. $R_{\text{equiv}} = \dfrac{R_1 R_2}{R_1 + R_2}$	equation for equivalent parallel resistors
2. $R_{\text{equiv}} = \dfrac{(97\ \Omega)(20\ \Omega)}{97\ \Omega + 20\ \Omega}$	enter values
3. $R_{\text{equiv}} = 17\ \Omega$	arithmetic

Step 4



The circuit cannot be simplified any further. The question asks for the current at the bend in the circuit near the battery. We calculate this using Ohm's law. Note that the current would differ at other locations in the original circuit due to its branches.

Step	Reason
1. $\Delta V = IR$	Ohm's law
2. $I = \dfrac{\Delta V}{R}$	solve for I
3. $I = \dfrac{12\ \text{V}}{17\ \Omega} = 0.71\ \text{A}$	enter values and do arithmetic

29.17 - Kirchhoff's loop rule

Kirchhoff's loop rule: The sum of the changes in electric potential across all components around any complete loop of a circuit is zero.

German physicist Gustav Kirchhoff (1824-1887; his name is pronounced "keer-koff") developed two powerful rules that aid in the analysis of circuits. First we discuss his loop rule. It applies to any path around a circuit that begins and ends at the same point. Such a path is often called a closed path, or a closed loop. His loop rule states that the sum of the changes in electric potential is zero for **any** path that begins and ends at the same point. In Concept 1, you see such a path: It starts and stops at the same point below the battery.

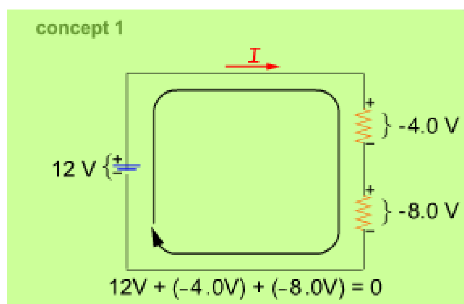
We implicitly derived this rule earlier when we considered a charge that completed a closed path journey around a circuit and stated that the charge's PE was the same at the beginning and end of its journey. We also used that analysis to conclude that the sum of the potential changes around the complete circuit loop must equal zero.

The circuit in Concept 1 illustrates Kirchhoff's loop rule. It is a series circuit with a battery and two resistors with different resistances. According to the loop rule, the potential changes across the two resistors and the battery must sum to zero.

It is important to consider carefully whether the potential changes are positive or negative across each device. If we traverse the circuit in Concept 1 in the direction of conventional current, then the change in potential is positive across the battery, and negative across each of the resistors.

We note the magnitudes of the changes in electric potential in the diagram, and use plus and minus signs on both the battery and the resistors to indicate where the potential is higher or lower.

We could also chart the path in a counterclockwise direction around the circuit in Concept 1. This direction is opposite the flow of conventional current. All the signs



Kirchhoff's loop rule
In circuit loop, sum of potential changes equals zero

concept 2

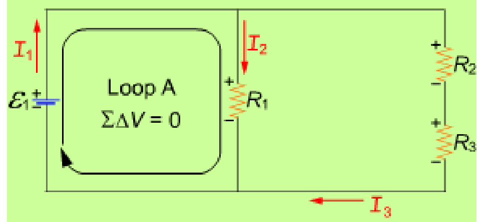
	With conventional current	Against conventional current
Resistor	$\Delta V = -IR$	$\Delta V = +IR$
	Move from – to + terminal	Move from + to – terminal
Ideal Battery	$\Delta V = +\mathcal{E}$	$\Delta V = -\mathcal{E}$

Determining sign of potential change across components

would be reversed − the change in potential would be negative across the battery and positive across the resistors − but they would still sum to zero.

The tables in Concept 2 summarize the guidelines for determining whether a potential change is positive or negative across resistors and ideal batteries. The tables also indicate how to quantify the potential change. For ideal batteries, the magnitude of the potential change equals the emf of the battery, $\mathcal{E}$. For resistors, we use Ohm's law to substitute IR for ΔV.

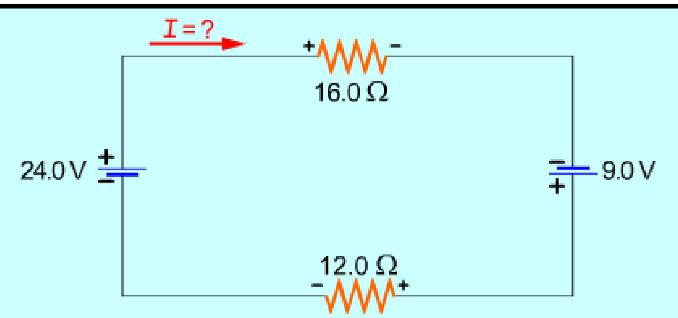
concept 3



Applying Kirchhoff to multi-loop circuits

Follow any complete loop around circuit
· Any complete path is a circuit loop
· Potential changes around any loop sum to zero
· Assess potential changes across each component

29.18 - Sample problem: Kirchhoff's loop rule



What is the current in this circuit?

Here is a circuit with two batteries and two resistors. We will use Kirchhoff's loop rule to solve for the current. To determine the sign of the potential change across each component, we will traverse the circuit in the direction of the conventional current.

Variables

emf of battery 1	24.0 V
emf of battery 2	9.0 V
resistance of resistor 1	$R_1 = 16.0\ \Omega$
resistance of resistor 2	$R_2 = 12.0\ \Omega$
potential difference across resistor 1	ΔV_{R1}
potential difference across resistor 2	ΔV_{R2}
current	I

What is the strategy?

1. Use Kirchhoff's loop rule to write an equation that includes all the potential changes in the circuit, setting their sum equal to zero.
2. The change in potential of a battery is the same as its emf. For the resistors, use Ohm's law to substitute current times resistance for potential difference. You may traverse clockwise. Be careful about signs.
3. Enter the given values for the potential differences of the batteries and the resistance of the resistors to solve for the remaining variable, the current in the circuit.

Physics principles and equations

The current is the same at all points in a series circuit.

Kirchhoff's loop rule states that the potential changes in a closed loop sum to zero. This circuit is a closed loop.

Ohm's law

$$\Delta V = IR$$

Step-by-step solution

First, we write an equation for the potential changes in the circuit using Kirchhoff's loop rule. We assess whether the potential changes are positive or negative across each component by moving around the circuit in the direction of conventional current.

Step	Reason
1. $\Delta V_{\text{bat1}} + \Delta V_{\text{R1}} + \Delta V_{\text{bat2}} + \Delta V_{\text{R2}} = 0$	Kirchhoff's loop rule
2. $+24.0\text{ V} + \Delta V_{\text{R1}} + \Delta V_{\text{bat2}} + \Delta V_{\text{R2}} = 0$	potential change across battery 1
3. $+24.0\text{ V} + (-IR_1) + \Delta V_{\text{bat2}} + \Delta V_{\text{R2}} = 0$	potential change across resistor 1
4. $+24.0\text{ V} + (-IR_1) + 9.0\text{ V} + \Delta V_{\text{R2}} = 0$	potential change across battery 2
5. $+24.0\text{ V} + (-IR_1) + 9.0\text{ V} + (-IR_2) = 0$	potential change across resistor 2

Now that we assessed the potential change across each of the four components, we can solve the equation for the unknown variable, the common current I.

Step	Reason
6. $+33.0\text{ V} = IR_1 + IR_2$	rearranging
7. $+33.0\text{ V} = I(R_1 + R_2)$	factor out I
8. $I = \dfrac{33.0\text{ V}}{R_1 + R_2}$	solve for I
9. $I = \dfrac{33.0\text{ V}}{16.0\ \Omega + 12.0\ \Omega}$	enter values
10. $I = 1.18\text{ A}$	evaluate

29.19 - Kirchhoff's loop rule and more complex circuits

In this section we show how to apply Kirchhoff's loop rule to some more complicated circuits. The same principles discussed earlier apply, as does the advice that you must be careful with signs.

The upper circuit contains two batteries labeled A and B. Battery A has a greater emf than battery B. It drives conventional current clockwise through battery B. Battery A might be part of a circuit used to recharge battery B. If battery B were the sole battery in the circuit, conventional current would flow counterclockwise, not clockwise.

As the Concept 1 diagram shows, this means the emf when the circuit is traversed in the direction of conventional current is positive for the battery A, and negative for the battery B. It is positive for A since current is flowing from the negative terminal to the positive, from a point of lower electric potential to a point of higher electric potential.

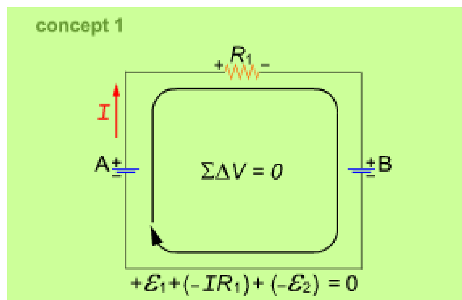
In battery B, current flows from a point of higher electric potential (the positive terminal) to the lower electric potential (the negative terminal). The potential change of the resistor is negative as well, since the direction of the path is the direction of conventional current.

Perhaps for some reason you erred in determining the direction of the conventional current, and guessed it went counterclockwise. In even more complex circuits, the direction of the current can be hard to determine in advance. If you guess wrong but still make your calculations correctly, all is not lost. You will calculate a negative value for the current, which indicates the current flows in the opposite direction to the one you first assumed.

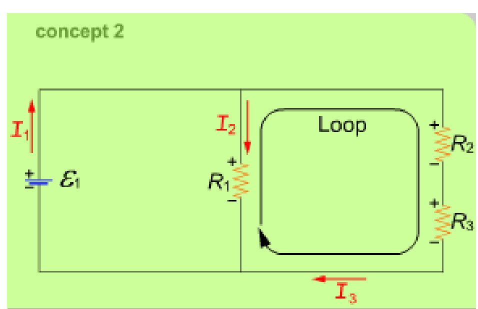
The diagram in Concept 2 shows a circuit used in a prior section. In the prior section, we used a loop that included the battery and R_1. Now we make a loop on the right side of the circuit, a loop that includes all three resistors.

We make this loop for one reason: to point out that the signs of the resistors differ. In the loop we make, we traverse the R_2 and R_3 resistors in the direction of conventional current, so the changes in electric potential are negative. However, R_1 is traversed in the direction opposite of conventional current, so the change there is positive. (One of the changes must be positive for them all to sum to zero.)

This analysis applies concepts discussed earlier; the purpose is that you cannot always count on the potential change being negative. If you are making arbitrary loops, you



Kirchhoff's rule: complex circuits
Loop drawn in direction of conventional current
· Potential changes sum to zero
· emf signs differ due to batteries' orientations



Another Kirchhoff loop

must be careful to consider the direction of conventional current in order to use the right sign.

Loop with no battery
Potential changes sum to zero
Sign of potential change in resistor depends on direction

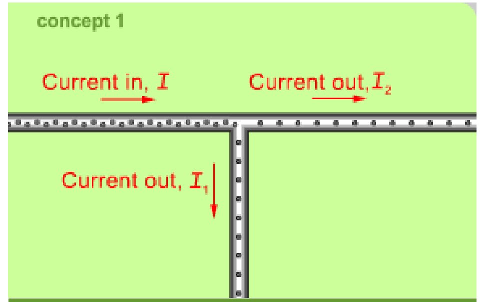
29.20 - Kirchhoff's junction rule

Kirchhoff's junction rule: The total current flowing into a junction equals the total current flowing out of a junction.

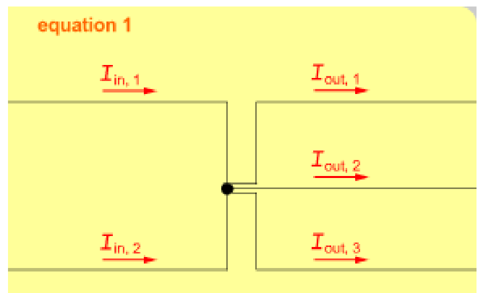
Kirchhoff's junction rule states that the amount of current that enters a junction must equal the amount of current that exits it. In other words, current in equals current out. This is shown as an equation in Equation 1.

Kirchhoff's rule is an implication of the principle of conservation of charge, which states that charge is neither created nor destroyed. The amount of charge flowing in equals the amount of charge flowing out.

The animation in Concept 1 shows the conservation of charge. The electrons enter the junction from below and from the right. They exit the junction on the left. The number of electrons exiting the junction equals the sum of the two flows of electrons entering the junction.



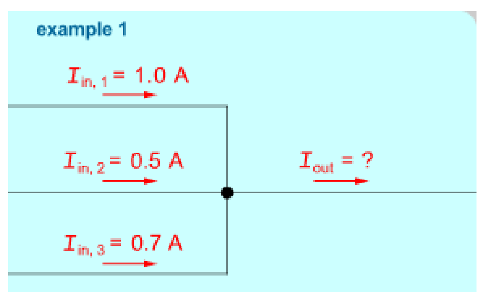
Kirchhoff's junction rule
Current into junction equals current out



Kirchhoff's junction rule

$$I_{in,1} + \ldots + I_{in,n} = I_{out,1} + \ldots + I_{out,m}$$

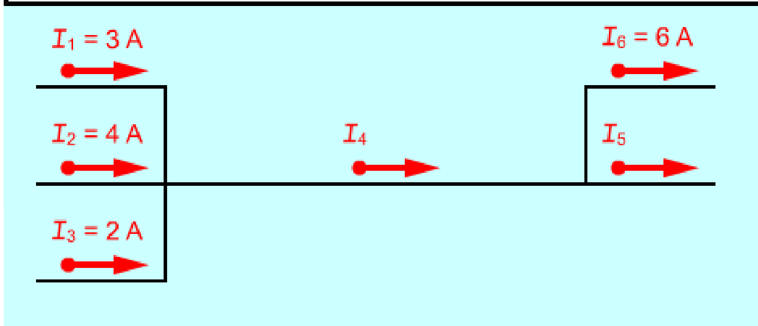
Current into junction = current out of junction



How much current exits this junction?

$I_{out} = I_{in,1} + I_{in,2} + I_{in,3}$
I_{out} = 1.0 A + 0.5 A + 0.7 A
I_{out} = 2.2 A

29.21 - Interactive checkpoint: Kirchhoff's junction rule

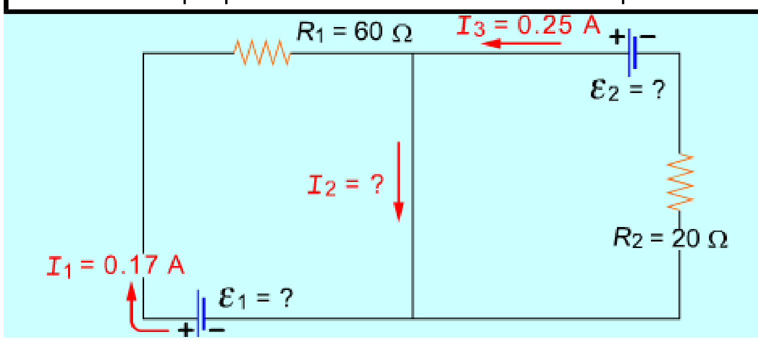


What is the current I_5?

Answer:

I_5 = [ ] A

29.22 - Sample problem: Kirchhoff's rules in a complex circuit



What are $\mathcal{E}_1$ and $\mathcal{E}_2$, and what is the current in the middle wire, I_2?

Here is a complex circuit with multiple resistors and multiple batteries in which the components cannot be simplified to equivalent components. We will use Kirchhoff's two rules to solve it.

What is the strategy?

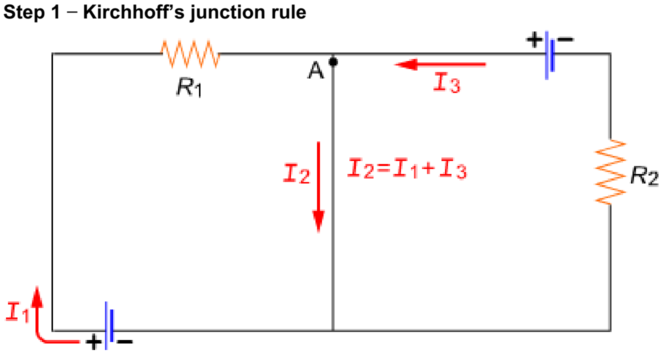
1. Identify possible currents and write an equation for currents using Kirchhoff's junction rule.
2. Sum potential changes around two loops and write an equation for each loop using Kirchhoff's loop rule.
3. The first two steps provide three equations and there are three unknowns. Solve for the current I_2 and then use that value to solve the two loop equations for the potential differences across the batteries.

Physics principles and equations

Kirchhoff's loop rule states the sum of the potential changes in any circuit loop equals zero.

Kirchhoff's junction rule states that current flowing into any junction equals current flowing out.

Step 1 – Kirchhoff's junction rule

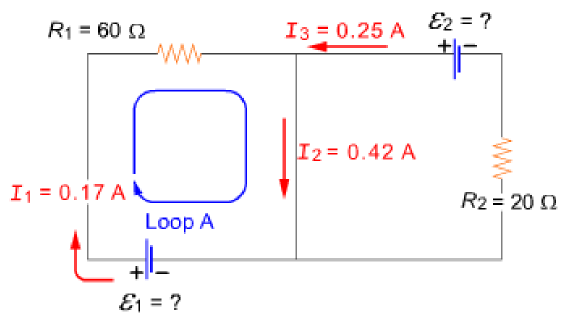


The first step is to look at the various currents. They will combine in the center wire, at the point marked A. We could probably guess the direction of the currents based on the orientations of the batteries, but we do not have to. We will add them, and the sign will tell us the direction. If we get a negative value for the I_2 current, it means that current points in the opposite direction.

We apply Kirchhoff's junction rule: I_2 equals the sum of I_1 and I_3. This is our first equation, and we can solve for I_2 as shown in the steps below.

Step	Reason
1. $I_2 = I_1 + I_3$	Kirchhoff's junction rule
2. $I_2 = 0.17\text{ A} + 0.25\text{ A}$	enter values
3. $I_2 = 0.42\text{ A}$	solve for I_2

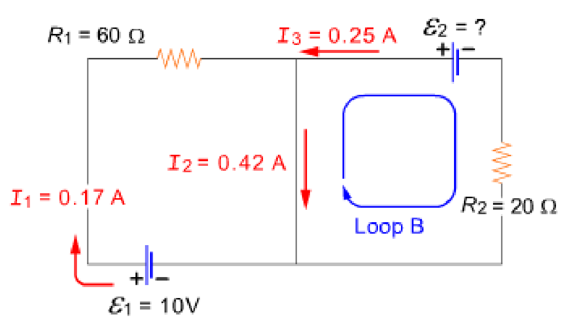
Step 2 – Kirchhoff's loop rule



Here we identify one of the loops of the circuit in order to apply Kirchhoff's loop rule and solve for the unknown battery emf, $\mathcal{E}_1$. We choose to make this loop in a clockwise fashion. This means the potential change is positive across the battery and negative across the R_1 resistor.

Step	Reason
1. $+\mathcal{E}_1 + \Delta V_{R1} = 0$	Kirchhoff's loop rule
2. $+\mathcal{E}_1 + (-I_1R_1) = 0$	assess potential change across resistor and use Ohm's law
3. $\mathcal{E}_1 = I_1R_1$	rearrange
4. $\mathcal{E}_1 = (0.17\text{ A})(60\ \Omega)$	enter values
5. $\mathcal{E}_1 = 10\text{ V}$	solve

Step 3 – Kirchhoff's loop rule again

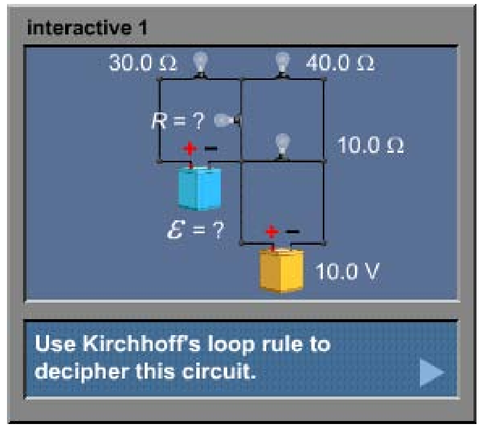


Now we apply Kirchhoff's loop rule to the other loop in order to determine the emf of the other battery, $\mathcal{E}_2$.

Step	Reason
1. $-\mathcal{E}_2 + \Delta V_{R2} = 0$	Kirchhoff's loop rule
2. $-\mathcal{E}_2 + I_3R_2 = 0$	assess potential change across resistor and use Ohm's law
3. $\mathcal{E}_2 = I_3R_2$	rearrange
4. $\mathcal{E}_2 = (0.25\text{ A})(20\ \Omega)$	enter values
5. $\mathcal{E}_2 = 5.0\text{ V}$	solve

29.23 - Interactive problem: a complex circuit

At the right, you see a complex circuit. Your investigative tool in this simulation is an ammeter. Your task: Analyze this circuit to determine the resistance of the middle light bulb and the emf of the battery on the left.



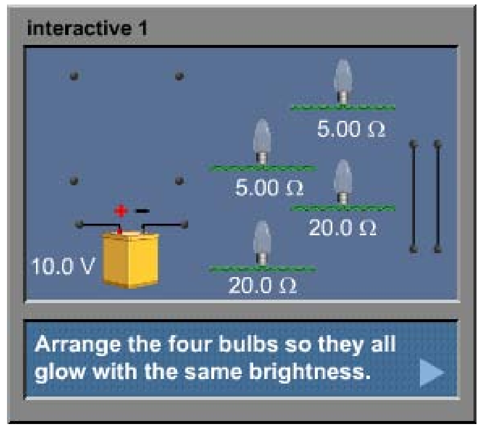
Some hints: Compute the potential changes in resistors with stated values using Ohm's law and the current readings from the ammeter. Then use Kirchhoff's loop rule to determine the unknown values.

Determine the unknown values R and $\mathcal{E}$, and enter these values in the answer boxes provided.

Press CHECK to see if your answers are correct. You can try again by entering new values and pressing CHECK again.

29.24 - Interactive problem: stringing holiday lights

In this interactive exercise, you have four bulbs for a strand of holiday lights. Two of the bulbs have a resistance of 20 Ω, and two have a resistance of 5 Ω.



Can you put them in the circuit so they all glow equally brightly? Put another way, can you determine a way to wire the four bulbs together so that each is supplied with the same amount of power? The brightness of a light bulb is proportional to the power supplied to it. The power dissipated by a light bulb equals the product of the current through it and the potential difference across it: $P = I\Delta V$. Using Ohm's law, you can also state that $P = I^2R$.

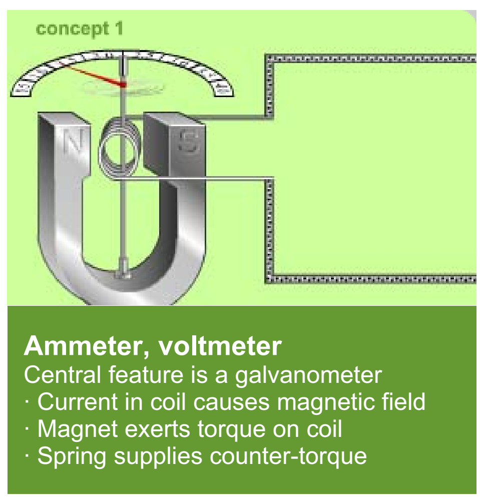
The simulation features a 10-volt battery, some wires and the four bulbs. You can build a circuit by dragging the wires and components into place on the grid on the screen.

The simulation also comes with an ammeter and voltmeter. If the bulbs are not all the same brightness, you can use the ammeter and voltmeter to try to determine where you went wrong.

If at any time you believe you have wired the bulbs correctly, press CHECK. A text message will say whether the bulbs are arranged correctly or not. If not, rearrange the bulbs to try again, and press CHECK again to test your answer.

29.25 - Physics at work: ammeters and voltmeters

Ammeters and voltmeters can be made with either digital or analog circuitry. This section describes how an analog device functions.



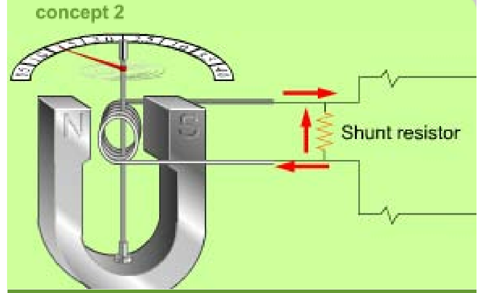
The central mechanism of an ammeter or voltmeter is a *galvanometer*. The essential components of a galvanometer include a magnet, a coil of wire, a spring, a needle and a scale. A galvanometer is shown on the right.

The needle is attached to the coil as shown. The coil is in the magnetic field created by the magnet. When current flows, the current-carrying coil creates its own magnetic field (the general relationship of currents and magnetic fields is discussed later). The strength of this magnetic field is proportional to the current.

The magnetic field of the magnet exerts a torque on the coil because of the coil's magnetic field. (Imagine a bar magnet between the magnet shown on the right. It will be rotated to be parallel with the field created by the U-shaped magnet.) The amount of torque is in part a function of the amount of current flowing through the coils. A spring opposes the motion of the coil, applying a counter torque. The force (and torque) exerted by the spring increases the more it is compressed. When the two torques are equal, the coil stops rotating, and the needle points to a value on a scale.

The galvanometer itself is designed to function within a certain range of current. When it is used as an ammeter, additional circuitry allows it to measure a wide variety of currents, as shown in Concept 2. A variable resistor functions as a *shunt resistor*. By adjusting the resistance of the shunt resistor in the parallel branch of the circuit, the user insures that an appropriate amount of current flows through the galvanometer.

concept 2



Shunt resistor
Controls amount of current through galvanometer

29.26 - Capacitors in series

To calculate the equivalent effect of capacitors in series: Add the reciprocals of each component's capacitance. The reciprocal of the sum is the equivalent capacitance.

You can simplify circuits involving capacitors by calculating the equivalent capacitance of capacitors in series. The equation for doing so is shown in Equation 1. The reciprocals of each individual capacitance are added. The reciprocal of that value equals the equivalent capacitance. That is perhaps better stated as an equation than in words!

There is also a useful shortcut for calculating the equivalent capacitance of two capacitors in a series. It is also shown in Equation 1, and it can be derived from the first equation using algebra.

The general equation for calculating the equivalent capacitance in series is derived below. It is based on an important premise: The amount of charge on the plates of each capacitor in a series circuit must be the same. Why? Consider the diagram in Equation 1, starting with the capacitor C_1. The charge on each capacitor plate is equal but opposite since in a capacitor, the plates always have equal but opposite amounts of charge. The charge carriers from the bottom plate of C_1 flow to the top of C_2, but they must stop there because of the gap between the plates.

The bottom plate of C_2 must have the same charge as its top plate. The process then continues with C_3: the charge from the bottom of C_2 flows to the top plate of C_3, so it, too, has the same amount of charge, and so on. If the capacitances of these capacitors differ, then the potential differences across them differ as well, but the charge is identical.

Derivation.

Variables

potential difference across all capacitors	ΔV_{total}
potential difference across i^{th} capacitor	ΔV_{i}
charge on equivalent capacitor plate	q
charge on i^{th} capacitor	q_{i}
equivalent capacitance of circuit	C_{equiv}
capacitance of i^{th} capacitor	C_{i}

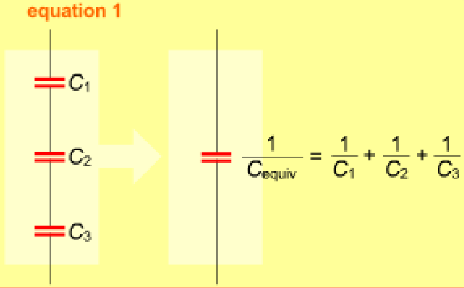
Strategy

1. Find the total potential difference across all the capacitors combined.
2. Use the definition of capacitance to rewrite potential difference in terms of charge and capacitance. An algebraic simplification gives the series rule for capacitors.

Physics principles and equations

With components in series, the total potential difference across all of them equals the sum of the potential differences across each component.

equation 1



Capacitors in series

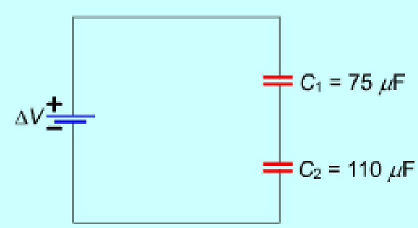
$$\frac{1}{C_{\text{equiv}}} = \frac{1}{C_1} + \frac{1}{C_2} + \ldots + \frac{1}{C_n}$$

Reciprocal of equivalent capacitance = sum of reciprocals of capacitances

Shortcut for two capacitors:

$$C_{\text{equiv}} = \frac{C_1 C_2}{C_1 + C_2}$$

example 1



What is the equivalent capacitance of these two capacitors?

$$\frac{1}{C_{\text{equiv}}} = \frac{1}{C_1} + \frac{1}{C_2}$$

$$\frac{1}{C_{\text{equiv}}} = \frac{1}{75\ \mu\text{F}} + \frac{1}{110\ \mu\text{F}}$$

Definition of capacitance

$$C = \frac{q}{\Delta V}$$

Step-by-step derivation

We use the principles stated above to derive the equation for capacitors in series.

Step	Reason
1. $\Delta V_{total} = \Delta V_1 + \Delta V_2 + \ldots + \Delta V_n$	total ΔV is sum of series ΔV's
2. $\frac{q}{C_{equiv}} = \frac{q_1}{C_1} + \frac{q_2}{C_2} + \ldots + \frac{q_n}{C_n}$	definition of capacitance
3. $\frac{1}{C_{equiv}} = \frac{1}{C_1} + \frac{1}{C_2} + \ldots + \frac{1}{C_n}$	divide by equal charge

$$\frac{1}{C_{equiv}} = 0.022\ (\mu F)^{-1}$$

$$C_{equiv} = \frac{1}{0.022\ (\mu F)^{-1}} = 45\ \mu F$$

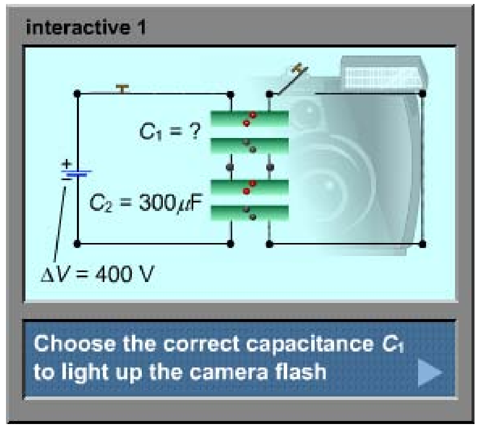
29.27 - Interactive problem: capacitors in a camera flash

A flash bulb in a camera requires that a large amount of energy flow through it in a short period of time. Cameras use capacitors to store this energy and make it available in an instant.

In the simulation to the right, your goal is to store the correct amount of charge, 0.04 coulombs, in the capacitor circuit. There is a 400-volt potential difference in the circuit. (Cameras typically contain low voltage batteries, such as three volts, but their circuitry increases this voltage so that a potential difference of 300 to 500 volts is created across the capacitor plates.)

The capacitor C_2 in the diagram has a capacitance of 300 μF. If it is used alone, the potential difference across it will be too great to be safe. A second capacitor C_1 has been added to the circuit in series with C_2 to reduce the required potential difference across C_1.

Your challenge is to determine what the capacitance of C_1 should be to give the pair of capacitors an equivalent capacitance such that the charge q stored by them is 0.04 coulombs.



This is a two-step problem. First, you must determine the required equivalent capacitance. Then, you must use the series rule for capacitors, stated in the previous section, together with the given capacitance of C_2 to determine the capacitance of C_1.

Set the C_1 capacitance in the simulation and press GO. If you have calculated correctly, the flash bulb illuminates. If not, a message tells you whether your value for the C_1 capacitance is too high or too low. Be very sure that C_1 is not set too high!

29.28 - Capacitors in parallel

To calculate the equivalent capacitance of capacitors in parallel: Add each capacitor's capacitance.

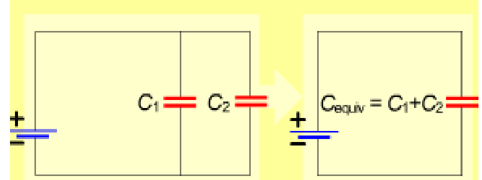
The equivalent capacitance of capacitors connected in parallel equals the sum of each capacitor's capacitance.

The example problem on the right shows two capacitors in parallel. One has a capacitance of 25 μF, the other 42 μF. The equivalent capacitance is the sum of the two: 67 μF.

Let's determine the charge on a single plate of each of these two parallel capacitors when a 10-volt battery is connected in parallel with them. Since all the components are in parallel, the potential difference across each capacitor is 10 V.

The charge is computed using the equation $q = C\Delta V$. Multiplying each capacitance by 10 volts enables us to determine the charge. The 25 μF capacitor has a charge of 250 μC, and the 42 μF capacitor has a charge of 420 μC. Note that in parallel configurations, the potential difference across the capacitors is the same, but the charge on the capacitors differs when their capacitances differ.

equation 1



Capacitors in parallel

$$C_{equiv} = C_1 + C_2 + \ldots + C_n$$

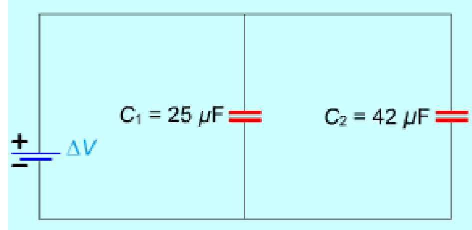
Equivalent capacitance = sum of capacitances

Derivation. In the steps below, we derive the equation for finding the equivalent capacitance of capacitors in parallel. Note that q_i refers to the charge on **a** plate of a capacitor, and that q_{total} is the sum of the charges on these plates. (The total charge across all plates is zero, since capacitors are electrically neutral: q always refers to the charge on a single plate in the context of capacitors.)

Variables

total charge on all capacitors	q_{total}
charge on a plate of i^{th} capacitor	q_i
potential difference across all capacitors	ΔV
potential difference across i^{th} capacitor	ΔV_i
equivalent capacitance of circuit	C_{equiv}
capacitance of i^{th} capacitor	C_i

example 1



What is the equivalent capacitance in this circuit?

$C_{equiv} = C_1 + C_2$

$C_{equiv} = 25\ \mu F + 42\ \mu F = 67\ \mu F$

Strategy

1. Find the total charge of all the capacitors combined.
2. Use the definition of capacitance to rewrite charge in terms of capacitance and potential difference. An algebraic simplification gives the parallel rule for capacitors.

Physics principles and equations

Definition of capacitance

$$C = \frac{q}{\Delta V}$$

The potential difference is the same across all parallel branches in a circuit.

Step-by-step derivation

The top illustration on the right shows two capacitors in parallel. In the following steps, we assume that there are *n* capacitors in parallel, similarly arranged.

Step	Reason
1. $q_{total} = q_1 + \ldots + q_n$	total charge equals sum of each charge
2. $C_{equiv}\Delta V = C_1\Delta V_1 + \ldots + C_n\Delta V_n$	definition of capacitance
3. $C_{equiv} = C_1 + \ldots + C_n$	divide by equal potential difference

29.29 - Sample problem: powering a laser

A capacitor bank is made up of 5184 high-voltage capacitors, connected in parallel. Each capacitor has a capacitance of 299 μF.

What is the equivalent capacitance of the entire bank?

The National Ignition Facility in Livermore, CA houses the largest laser in the world. For each laser pulse, a bank of capacitors emits a burst of power that for a brief instant is a thousand times greater than the average power consumption of the entire United States. The capacitor bank consists of the parallel set of capacitors described above. You are asked to find their equivalent capacitance.

Variables

capacitance of single capacitor	$C = 299\ \mu F$
number of capacitors	$n = 5184$
equivalent capacitance of bank	C_{equiv}

What is the strategy?

1. Add the capacitances. Or perhaps better put, multiply one capacitance by the number of capacitors.

Physics principles and equations

The equivalent capacitance of capacitors in parallel is

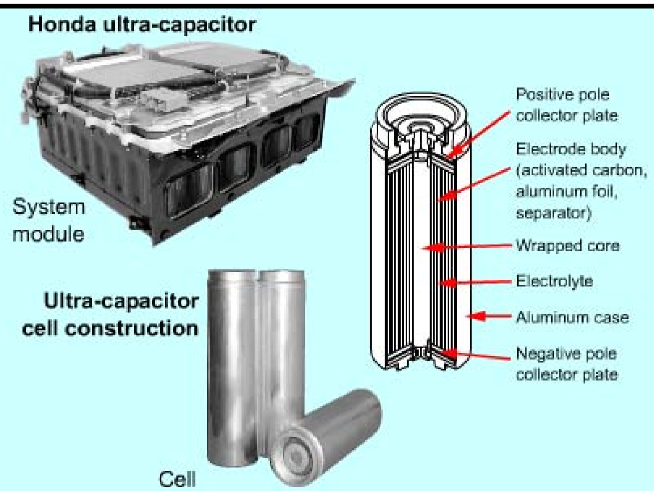
$$C_{\text{equiv}} = C_1 + C_2 + \ldots + C_n$$

Step-by-step solution

We rewrite the formula for n identical capacitors, and substitute the values stated in the problem above.

Step	Reason
1. $C_{\text{eq}} = nC$	n identical capacitors
2. $C_{\text{equiv}} = (5184)(299 \times 10^{-6}\ \text{F})$ $C_{\text{equiv}} = 1.55\ \text{F}$	evaluate

29.30 - Interactive checkpoint: return of the fuel cell car



If an 8.00-farad equivalent ultracapacitor contains 36 identical cells wired in parallel, what is the individual capacitance of each cell? Find what the individual capacitance of each cell would have to be if you wired the capacitors in series to make the ultracapacitor.

Answer:

$C_P =$ ______ F

$C_S =$ ______ F

29.31 - Circuits with resistors and capacitors (RC circuits)

Electric circuits can contain both resistors and capacitors. These kinds of circuits are called *RC circuits*. An *RC* circuit is shown in the diagram of Equation 1. The amount of current flowing through such a circuit varies over time, a phenomenon which we will discuss in this section. This variability is in contrast to the earlier sections of this chapter where current was presumed to be constant and we analyzed DC circuits operating in a steady state.

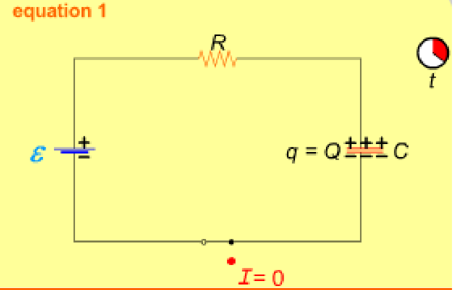
It takes time for a capacitor to charge. The time required is typically fleeting by human standards, but it is very important in circuits that use capacitors for timing purposes.

When the switch is closed in the circuit shown on the right, charge begins to flow and accumulate on the plates of the capacitor. When enough charge accumulates to make the potential difference across the capacitor equal and opposite to that supplied by the battery, the current stops and the circuit is in equilibrium. (You can refresh your browser if you missed this.)

According to Kirchhoff's loop rule, the sum of the potential changes across the resistor, the capacitor and the battery is zero at all times. We write this relationship as Equation 1.

The signs in the equation correspond to traversing the circuit clockwise, in the direction of the conventional current. We cross the battery from its negative to its positive

equation 1



RC circuits

$$\mathcal{E} - IR - \frac{q}{C} = 0$$

$\mathcal{E}$ = battery emf

terminal, meaning the potential change across the battery is positive. The potential change across the resistor is negative. The potential change across the capacitor is also negative (since we cross from its positively charged to its negatively charged plate.)

We will use Q to represent the final charge on the capacitor, and q to represent the charge at any moment in time. When the circuit switch is first closed (at time $t = 0$), there is no charge on the capacitor ($q = 0$) and no potential difference across it to oppose the flow of current, so we may use Ohm's law to state that the initial current is $I_0 = \mathcal{E}/R$. The current then charges the capacitor, quickly creating a potential difference across it that is equal in magnitude to the battery's emf. When the capacitor is fully charged, the current is zero, and we can use the definition of capacitance to state that the final charge is $Q = C\mathcal{E}$.

The charge q on the capacitor plates increases and the current I through the circuit decreases as the capacitor charges up from $q = 0$ to $q = Q$. In Equations 2 and 3 to the right, you see graphs and equations describing the charge on the capacitor plates and the current in the circuit as functions of time. The equations both include exponential functions of time, with the base $e = 2.718\ 28$ being the base of the natural logarithms.

The product RC appearing in the exponent in each equation is called the *time constant* of the circuit, and is sometimes represented by the Greek letter τ (tau). This is the time required for the current I to decrease to $1/e$ (approximately 37%) of its original value I_0. It is also the time it takes for the charge q to increase from zero to $1 - 1/e$ (approximately 63%) of its final value Q.

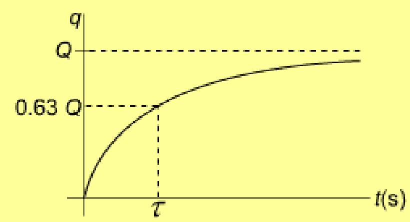
The units for τ are seconds. You can confirm this fact if you multiply the unit of resistance (1 Ω = 1 V/A = 1 V·s/C) by the unit of capacitance (1 F = 1 C/V). Everything but seconds cancels out. If a circuit has a resistance of 10 ohms and a capacitance of 6 microfarads, then τ = 60 microseconds. This means the capacitor will be 63% charged in 60 microseconds.

The battery supplies the energy to charge the capacitor. When the capacitor is fully charged, the energy stored in it equals $\frac{1}{2}Q\mathcal{E}$. Some energy is also dissipated as heat in the resistor. The energy dissipated in the resistor as the capacitor charges up equals $\frac{1}{2}Q\mathcal{E}$ as well.

The graphs on the right might cause you to think that a capacitor never fully charges up to its theoretical maximum, that the charge curve gets closer and closer to Q but never reaches it. Although this is mathematically true, the example problem on the right asks a more practical question: How long does it take the capacitor in a typical *RC* circuit to reach 99.99% of its maximum charge? To solve the problem, we note that this means the ratio of $q(t)$ to Q is 0.9999. The answer, as you can see, is that a capacitor charges very rapidly.

I = current in circuit, R = resistance
q = capacitor charge, C = capacitance

equation 2

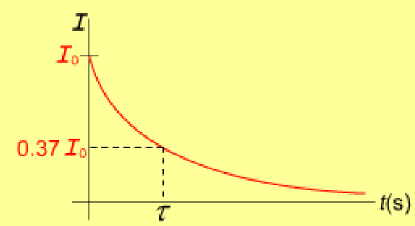


Charge as a function of time

$$q(t) = Q(1 - e^{-t/RC})$$

q = charge, t = time
Q = maximum charge
$RC = \tau$, time constant of circuit

equation 3

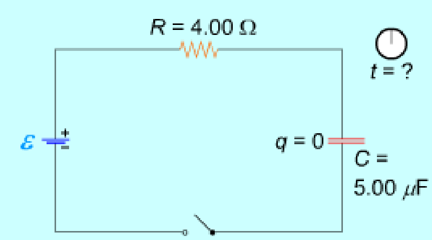


Current as a function of time

$$I(t) = \frac{\mathcal{E}}{R} e^{-t/RC}$$

I = current

example 1



How long after the switch is closed will the capacitor reach 99.99% of its full charge?

$q(t) = Q(1 - e^{-t/RC})$

$1 - e^{-t/RC} = q(t)/Q = 0.9999$

$e^{-t/RC} = 0.0001$

$t = -RC\ln(0.0001)$

$t = -(4.00\ \Omega)(5.00\times10^{-6}\ \text{F})(-9.21)$
$t = 184\times10^{-6}\ \text{s} = 184\ \mu\text{s}$

29.32 - Discharging a capacitor in an RC circuit

This electronic metronome uses a timing circuit to mark out rhythmic beats.

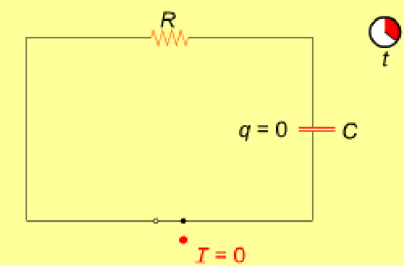
In the circuit shown in Equation 1, a battery has been used to charge the capacitor, the switch has been opened and the battery has been removed. As you watch, the switch is again closed, and a current flows until the capacitor is fully discharged. (You can refresh your browser if you missed this.) In this section, we discuss how long it takes for the capacitor to discharge when the switch is closed.

Using Kirchhoff's loop rule, Equation 1 states that the sum of the potential changes across the capacitor and the resistor is zero. The signs in the equation correspond to traversing the circuit counterclockwise, which is the direction of the conventional current. We encounter a positive change in potential as we cross the capacitor from its negative to its positive plate and a negative change in potential across the resistor.

The graphs that accompany the charge and current equations portray exponential decay. In the discharge circuit, both the charge and the current start at their maximum values and diminish toward zero. The time constant of the circuit is $\tau = RC$, as it is with a charging circuit. After this amount of time has passed, only 37% of the initial charge Q is left on the capacitor, and 37% of the original current I_0 flows through the circuit.

An *RC* circuit is the basis for many timing devices. Two applications of *RC* circuits are to intermittent windshield wipers (when the timing capacitor that controls them is fully discharged, its cycle of charging starts anew) and to cardiac pacemakers. An *RC* circuit in a pacemaker provides the timing, so that the pacemaker sends an electric current through the heart at regular intervals, ensuring that the heart muscle contracts at a regular rate.

equation 1

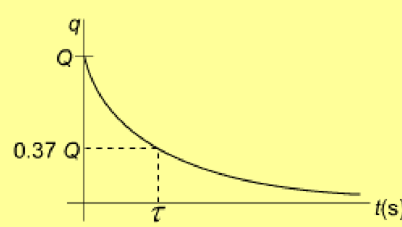


Discharging a capacitor

$$\frac{q}{C} - IR = 0$$

q = capacitor charge, C = capacitance
I = current, R = resistance

equation 2

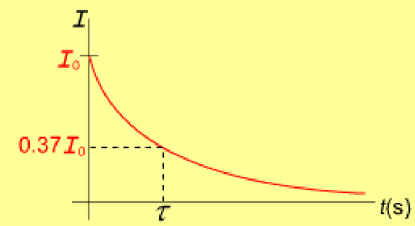


Charge as function of time

$$q(t) = Qe^{-t/RC}$$

q = charge, t = time
Q = maximum charge
$RC = \tau$, time constant of circuit

equation 3

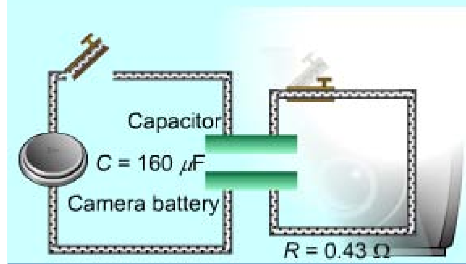


Current as a function of time

$$I(t) = \frac{Q}{RC} e^{-t/RC}$$

I = current

example 1



A camera has a flash discharge circuit with a resistance of 0.43 Ω. How long does it take to discharge 99% of its initial charge?

$q(t) = Qe^{-t/RC}$

$e^{-t/RC} = q(t)/Q = 0.010$

$t = -RC \ln(0.010)$

$t = -(0.43\ \Omega)(1.6 \times 10^{-4}\ \text{F})(-4.61)$

$t = 3.2 \times 10^{-4}\ \text{s} = 320\ \mu\text{s}$

29.33 - Gotchas

Components in series must have the same potential difference across them. No, they have the same amount of current flowing through them.

The potential difference across components in parallel is the same. Yes.

29.34 - Summary

An electric circuit is a set of electric components such as batteries, capacitors and resistors that are connected directly or through wires.

Current will only flow through a circuit that is closed, which means that it makes a loop. To cause current to flow continuously through a circuit, a source of emf, $\mathcal{E}$, such as a battery, must be present.

One basic type of circuit wiring is series wiring. Two circuit components are said to be in series when current must go through both components: there is only one possible path. The same amount of current flows through both components, and the sum of the potential differences across the components is the net potential difference across the combination.

Kirchhoff's loop rule says that the sum of all the potential differences (including any batteries) is zero around any circuit loop.

Resistances in series add to give the equivalent resistance of a combination of resistors.

The sum of the reciprocals of capacitances wired in series equals the reciprocal of the equivalent capacitance of a combination of capacitors.

The other basic type of wiring is parallel wiring. Components are connected so that they have the same potential difference across them, but may have different currents through them.

Kirchhoff's junction rule helps determine the specific currents. It says that the current flowing into a circuit junction must equal the current flowing out.

The sum of the reciprocals of resistances wired in parallel equals the reciprocal of the equivalent resistance of a combination of resistors.

Capacitances in parallel add to give the equivalent capacitance of a combination of capacitors.

An RC circuit consists of a resistor and a capacitor in series. *RC* circuits are useful in applications that require timing because the choice of the components determines how long the capacitor takes to charge or discharge.

Equations

Resistors in series

$$R_{\text{equiv}} = R_1 + R_2 + \ldots + R_n$$

Resistors in parallel

$$\frac{1}{R_{\text{equiv}}} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$$

Capacitors in series

$$\frac{1}{C_{\text{equiv}}} = \frac{1}{C_1} + \frac{1}{C_2} + \ldots + \frac{1}{C_n}$$

Capacitors in parallel

$$C_{\text{equiv}} = C_1 + C_2 + \ldots + C_n$$

Chapter 29 Problems

Conceptual Problems

C.1 A cheap flashlight is dropped into seawater. It is "short-circuited". What does this expression mean?

C.2 Why do batteries rated at the same emf sometimes differ so much in size? For example, consider a 1.5 AA penlight battery compared to a D size battery.

C.3 What would happen if you directly connected the terminals of a battery without putting anything else in the circuit (this situation is called a "short circuit")?

C.4 Why would a voltmeter with a low amount of resistance be a mediocre tool?

C.5 Consider a circuit containing a battery and several resistors. You measure the current flowing out one end of the battery, and also measure the current entering the other end. Do you expect these measurements to be the same or different? Explain.

○ The same ○ Different

C.6 A battery is connected to either (1) three 5 Ω resistors wired in series or (2) one 10 Ω resistor. Which configuration will have the most current flowing through it?

○ Configuration (1) ○ Configuration (2)

C.7 You have five identical light bulbs. You build one circuit with a battery and two light bulbs in series and build another circuit with an identical battery and three light bulbs in series. The circuits are otherwise identical. In which circuit do you expect the light bulbs to shine brighter? Explain.

i. Circuit with two light bulbs
ii. Circuit with three light bulbs

C.8 Four 1.5 V batteries wired "in parallel" power an electronic device. If all but one of the batteries are removed, under what circumstances (if any) will the device continue to function normally?

C.9 Are the light fixtures in a typical building wired in series or in parallel? How do you know?

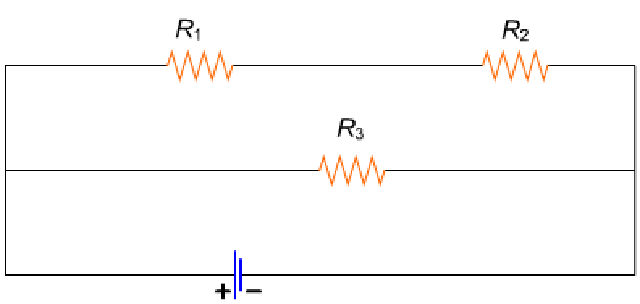
○ Series ○ Parallel

C.10 In some strings of holiday lights, especially cheaper or older ones, if one light "burns out", all the other lights in the string also go dark. (a) Are these types of lights wired in series or parallel? (b) Why does one light burning out cause the rest to go dark?

(a) ○ Series ○ Parallel
(b)

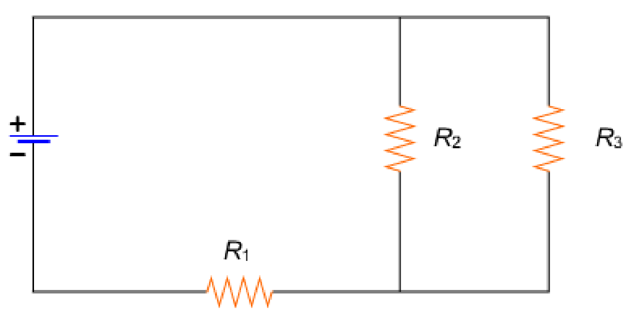
C.11 Why might an engineer design a circuit with components in parallel?

C.12 (a) Which resistors in the diagram are wired in series? (b) Which are wired in parallel?



(a) ○ R_1 and R_2 are in series
○ R_1 and R_3 are in series
○ R_2 and R_3 are in series
(b) ○ R_1 and R_2 are in parallel
○ R_1 and R_3 are in parallel
○ R_3 is in parallel with the combination of R_1 and R_2

C.13 (a) Which resistors are in series? (b) Which are in parallel?

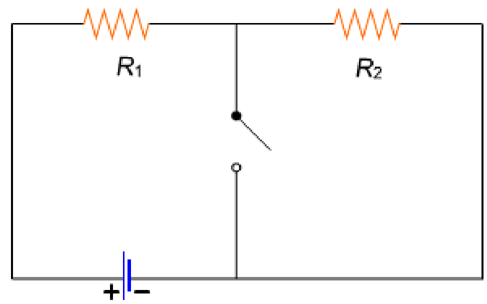


(a) ○ R_1 is in series with R_2
○ R_1 is in series with R_3
○ R_1 is in series with the combination of R_2 and R_3
○ R_2 is in series with R_3
(b) ○ R_1 is in parallel with R_2
○ R_1 is in parallel with R_3
○ R_1 is in parallel with the combination of R_2 and R_3
○ R_2 is in parallel with R_3

C.14 There are three light bulbs in parallel, powered by a single battery. A fourth bulb is added in parallel. (a) What effect does this have on the three other light bulbs? (b) What effect is there on the battery that is powering the circuit?

(a)
(b)

C.15 R_1 and R_2 are identical. At first, the switch is open. When the switch is closed, what happens to the power dissipated by R_2? Why?

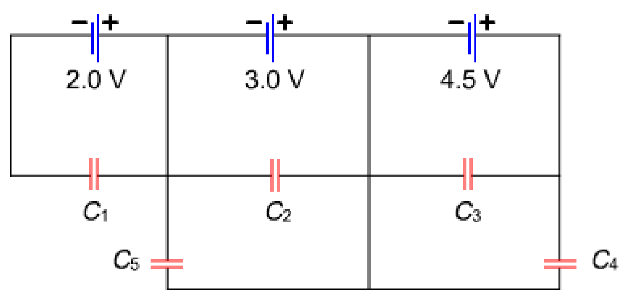


i. It will increase
ii. It will stay the same
iii. It will decrease to a positive value
iv. It will decrease to zero

C.16 For part of a circuit you are building, you need the highest capacitance possible. You only have a 10 pF capacitor and a 20 pF capacitor. Should you connect the capacitors in parallel, in series, or just use one capacitor by itself? Explain.

i. Parallel
ii. Series
iii. Use the 10 pF capacitor only
iv. Use the 20 pF capacitor only

C.17 What is the magnitude of the potential difference across (a) C_1, (b) C_2, (c) C_3, (d) C_4, (e) C_5? The answers require no calculations.



(a) ________ V
(b) ________ V
(c) ________ V
(d) ________ V
(e) ________ V

C.18 Suppose you connect a battery, a light bulb, and a capacitor in series. When the circuit reaches equilibrium, that is, after a very long time, will the light bulb be lit? Explain.

○ Yes ○ No

C.19 You start with a circuit containing only a battery and a capacitor. Is the capacitor's **equilibrium** charge affected when you (a) place a resistor in series with the capacitor or (b) place a resistor in parallel? (c) Explain.

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c)

Section Problems

Section 0 - Introduction

0.1 Leave the simulation in the interactive problem in this section in its initial state to answer the following questions. (a) How does the potential difference across the battery compare with the potential difference across the light bulb? (b) Is the current the same everywhere, or does it differ from place to place?

(a) ○ They are equal
○ They are unequal
(b) ○ It is the same everywhere
○ It differs from place to place

Section 2 - Electromotive force

2.1 A boat's battery has 325 watt-hours of available energy. How many joules of energy will it supply over its lifetime?

________ J

2.2 A battery has an emf of 7.5 V and 16 A of current flows through it. What power is this battery supplying?

________ W

Section 3 - Energy and electric potential in a circuit

3.1 You have a battery with unusual packaging: It is rated by how much work it does on charge. It says it does 6.6 J of work on every 2.2 C of charge. What is the emf of this battery?

__________ V

3.2 How much work does a 1.50 V flashlight battery do on an electron as it passes through?

__________ J

3.3 A battery charger is connected to a "dead" battery and causes 4.4 A of current to flow through the battery for 1.5 hours. It maintains a potential difference of 1.5 V across the terminals of the battery. (a) What is the power the charger delivers? (b) How much energy does it supply to the battery during the 1.5 hours?

(a) __________ W
(b) __________ J

3.4 There are two light bulbs in a circuit, powered by one battery. One light bulb is a 100 watt light bulb and the other a 200 watt light bulb. How much power must the battery be supplying for the lights to be running at full power?

__________ W

3.5 A battery in a circuit has an emf of 8.8 V. The only other component in the circuit is a resistor. 4.4 A of current flows in the circuit. What must the resistance of the resistor be?

__________ Ω

3.6 While the repair shop works on your voltmeter, they loan you an unusual replacement: Rather than measuring the potential difference across a resistor, it measures the amount of charge that flowed, and the change in electric potential energy of that charge. You use the device on a resistor and it tells you that 2.5 C of charge flowed through it, and that the electric potential energy of that charge decreased by 0.40 J. What is the potential difference across the resistor? (Report the answer as a positive number.)

__________ V

3.7 A specialized 8.4-volt battery will supply 430 C of charge over its lifetime. Assume that the potential difference it can maintain between its terminals does not change as it ages. It is wired into a simple circuit with a 350 ohm resistor. What is the resulting lifetime of the battery?

__________ s

Section 4 - Internal resistance of a battery

4.1 A battery with an emf of 1.5 V and a 12 Ω resistor are the only two components in a circuit. You measure the current flowing through the circuit to be 0.081 A. What is the internal resistance of the battery?

__________ Ω

4.2 When does the emf of a battery equal the potential difference across its terminals?

4.3 A battery is modeled as an ideal emf together with an internal resistance of 0.0270 ohms. A voltmeter says the potential difference across its terminals is 8.87 V. 4.50 A of current flows through the battery. What is the emf of the battery?

__________ V

4.4 You buy a battery that is rated at 1.50 V, which is what the manufacturer has stated as its emf. You place it in a circuit which has 4.25 A of current flowing through it, and then measure the potential difference across the battery's terminals. The voltmeter you are using says it is 1.48 V. What is the internal resistance of the battery?

__________ Ω

4.5 A battery has a rated emf of 9.00 V. You have determined that its internal resistance is 0.0320 ohms. The potential difference you measure across the battery is 8.98 V. What current is flowing through the battery?

__________ A

4.6 A battery with an emf of 12.0 V has an internal resistance of 0.0400 ohms. 5.00 A of current flows through the circuit. Of the power it supplies, what percentage is being wasted due to its internal resistance?

__________ %

4.7 A battery is modeled as an ideal emf of 18.0 V together with an internal resistance. It is in a circuit where the only other component is a light bulb that is outputting heat and light energy at a rate of 75.0 W. The potential difference across the battery system is 17.9 V. (a) What is the resistance of the hot light bulb? (b) What is the internal resistance of the battery?

(a) __________ Ω
(b) __________ Ω

4.8 The potential difference across a battery is 4.2 V when the current is 3.5 A, and 4.1 V when the current is 3.9 A. What is its internal resistance?

__________ Ω

Section 6 - Series wiring

6.1 Many flashlights are powered by a stack of batteries that are slid into the end of the flashlight one after another. Suppose you have a flashlight that takes four ideal 1.5 V batteries. If the resistance of the bulb is 11 ohms, what is the current through the bulb when the flashlight is turned on? The fact that the batteries are stacked end to end means they are wired in series.

__________ A

Section 7 - Resistors in series

7.1 A portable electric heater is powered by a 45 V battery. It has four identical heating elements inside, wired in series. If a current of 11 A runs through them when the heater is turned on, what is the resistance of each heating element?

__________ Ω

7.2 A 12 ohm, 14 ohm and 16 ohm resistor are connected in series. What is the equivalent resistance?

__________ Ω

7.3 Three hundred 14 ohm resistors are connected in series (think of a chain of cheap, old-fashioned holiday lights). What is the equivalent resistance?

__________ Ω

7.4 There are six identical resistors in series. A multimeter measures the total potential difference across **all** the resistors as 331 V. 4.56 A of current flows through the resistors. (a) What is the equivalent resistance of this group of resistors? (b) What is the resistance of each resistor?

(a) __________ Ω
(b) __________ Ω

7.5 The total potential difference across two resistors in series is 34.0 V. The current through them is 6.50 A. One resistor's resistance is 2.00 Ω. What is the other resistor's resistance?

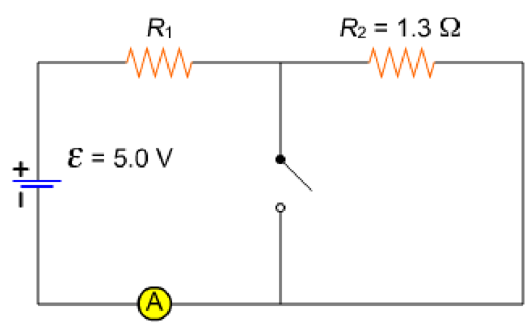
__________ Ω

7.6 A 55 Ω and 65 Ω resistor are wired in series in a circuit that contains a 12 V battery. (a) What is the equivalent resistance of the two resistors? (b) How much current is flowing in the circuit?

(a) __________ Ω
(b) __________ A

7.7 The ammeter in the diagram displays 1.6 A when the switch is open. What will it display when the switch is closed?

__________ A



7.8 A circuit contains a single resistor of 12.0 ohms and an ideal emf source (one with zero internal resistance). The current is 5.50 A. A second resistor is added in series, and the current falls to 4.23 A. What is the resistance of the second resistor?

__________ Ω

7.9 A 15.0 Ω and 25.0 Ω resistor are in series with a battery. The potential difference across the 15.0 ohm resistor is 4.50 V. What is the potential difference across the battery?

__________ V

7.10 Three silver wires of equal length are connected in series with a 1.50 V battery. Their cross-sectional areas are 1.00, 5.00 and 10.0 cm^2. (a) What is the potential difference across the narrowest wire? (b) The medium wire? (c) The widest wire?

(a) __________ V
(b) __________ V
(c) __________ V

Section 8 - Interactive problem: series wiring

8.1 Use the simulation in the interactive problem in this section to calculate the resistance of the R_2 bulb.

_________ Ω

Section 11 - Resistors in parallel

11.1 The four heating elements on an electric stove are wired in parallel so that any combination of them can be on at the same time. When an element is on, it behaves as a 6.70 Ω resistor. When all of the heating elements are on, what is the equivalent resistance of the stove?

_________ Ω

11.2 A 6.00 Ω and 42.0 Ω resistor are in parallel. What is their equivalent resistance?

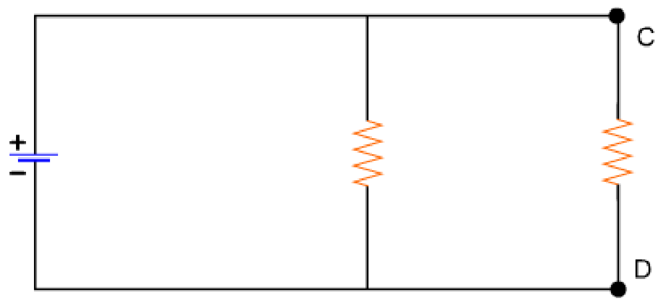
_________ Ω

11.3 Forty 7.00 Ω resistors are in parallel. What is their equivalent resistance?

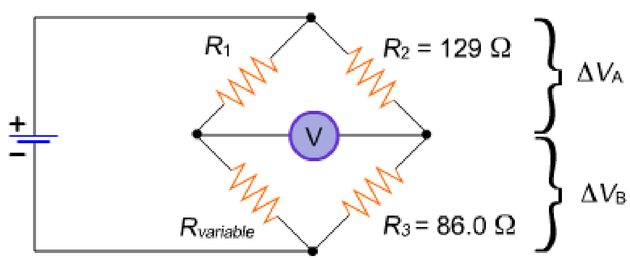
_________ Ω

11.4 The potential difference across points C and D is 9.2 V. The current flowing past C is 4.6 A. The two resistors are identical. (a) What is the resistance of each one? (b) What is their equivalent resistance?

(a) _________ Ω
(b) _________ Ω



11.5 This circuit, called a *Wheatstone bridge*, is useful for making precise resistance measurements. The potential difference measured by the voltmeter in the middle depends on the relative resistances of the four resistors. If that potential difference is zero, the Wheatstone bridge is said to be *balanced*. Suppose you balance the Wheatstone bridge in the figure by adjusting the variable resistor to 51.0 ohms. What is the resistance of R_1? (Hint: Because the bridge is balanced, the potential difference across R_1 equals that across R_2. The same relation holds between $R_{variable}$ and R_3.)



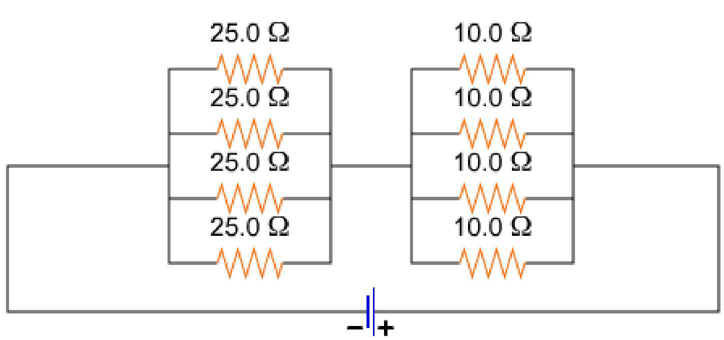
_________ Ω

11.6 You have access to many 12.5 Ω and 50.0 Ω resistors. (a) How can you create a 62.5 Ω equivalent resistor? (b) A 10.0 Ω equivalent resistor? (c) A 60.0 Ω equivalent resistor?

(a)
(b)
(c)

11.7 What is the equivalent resistance of the circuit shown?

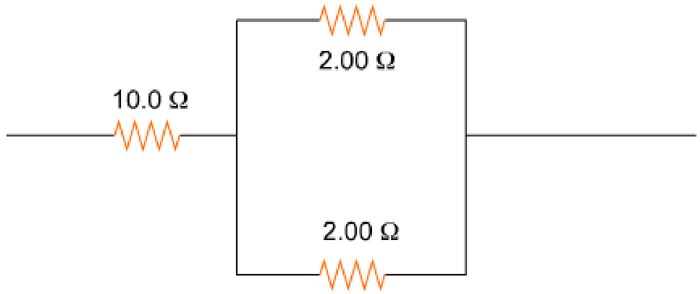
_________ Ω



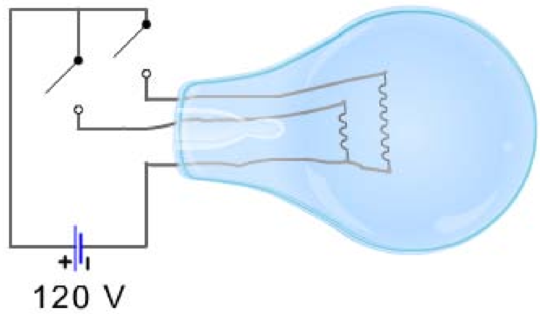
11.8 You have ten 4.0 Ω resistors. How can you create an equivalent resistance of 5.0 Ω? (There is more than one correct solution, and you do not have to use all the resistors.)

11.9 When the 10.0 Ω resistor starts to dissipate 100 W of power, it will rupture. What minimum emf can be applied to this circuit to cause the resistor to rupture?

_________ V



11.10 Many lamps have a switch that lets you choose three levels of brightness. The bulbs in this type of lamp have not one but two filaments inside, which are wired in parallel and have their own switches. One of the filaments is designed to give off more light than the other when connected to the average 120 V emf provided by a wall outlet. In the dimmest position, only the least bright filament is on. In the next position, only the brighter filament is on. In the brightest position, both filaments are on. Suppose one of the filaments consumes 55.0 watts and the other consumes 70.0 watts. What is their equivalent resistance in the brightest position (when both are turned on)?



_________ Ω

11.11 You have 3 identical resistors. How many equivalent resistances can you create, using all three resistors?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6

11.12 There are *n* identical resistors in a section of a circuit. The potential difference across this section of the circuit is ΔV. The resistors can be placed either all in series, or all in parallel. (a) Which configuration "consumes" more power, series or parallel? (b) By what factor must you multiply the power consumed by the low-power configuration to get the power consumed by the higher-power configuration?

(a) ○ Series ○ Parallel
(b) ○ $1/n$ ○ $1/n^2$ ○ n ○ n^2

11.13 In parallel, two resistors have an equivalent resistance of 100 ohms. In series, they have an equivalent resistance of 625 ohms. (a) What is the resistance of the resistor with greater resistance? (b) The lesser resistance?

(a) _________ Ω
(b) _________ Ω

Section 12 - Interactive problem: a parallel circuit

12.1 Use the simulation in the interactive problem in this section to calculate the resistance of the unknown resistor in the parallel circuit.

_________ Ω

Section 13 - Interactive problem: potential difference in parallel circuit

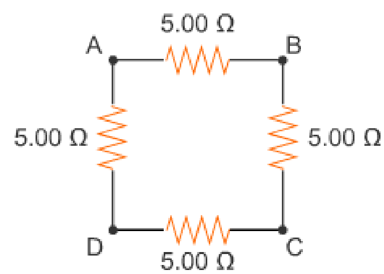
13.1 Use the information given in the interactive problem in this section to calculate (a) the unknown emf in the given circuit, (b) the unknown current, and (c) the unknown resistance. Test your answer using the simulation.

(a) _________ V
(b) _________ A
(c) _________ Ω

Section 14 - Circuits with series and parallel wiring

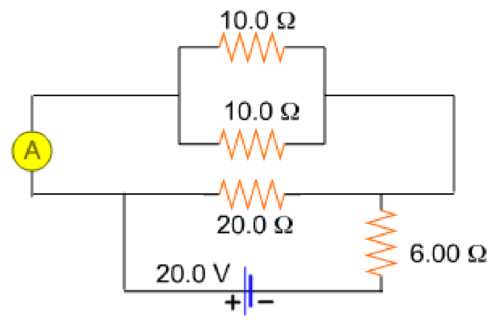
14.1 Four identical 5.00 Ω resistors are joined and form the four sides of a square. (a) What is the resistance between points A and B? (b) What is the resistance between A and C?

(a) ________ Ω
(b) ________ Ω



14.2 (a) What will the ammeter read? Report this as a positive number. (b) How much power is dissipated by one of the 10.0 Ω resistors?

(a) ________ A
(b) ________ W



14.3 Twelve identical 5.00 W resistors are joined and used to form the edges of a cube. What is the resistance between a pair of diagonally opposite vertices? Hint: If two points in a circuit are at the same potential, then you can mentally connect them and consider them as the same point.

________ Ω

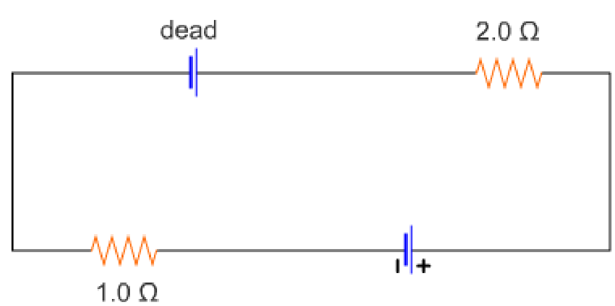
Section 17 - Kirchhoff's loop rule

17.1 A 9.0 V battery and three resistors are all connected in series. The potential difference across one resistor is measured to be 2.0 V, and the potential difference across another is measured to be 4.2 V. What is the potential difference across the third resistor?

________ V

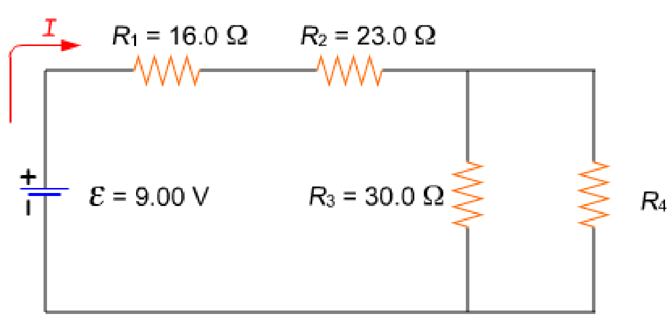
17.2 You want 5.0 A to flow through the upper battery which is "dead". What should the emf of the lower power source be? Assume that the batteries have negligible internal resistance.

________ V



17.3 In the circuit shown, the current I through the battery is 0.150 A. Find the current through R_3.

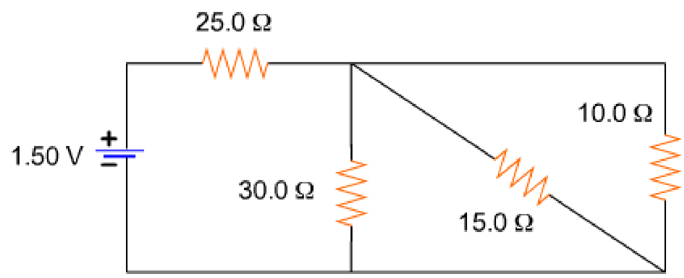
________ A



Section 19 - Kirchhoff's loop rule and more complex circuits

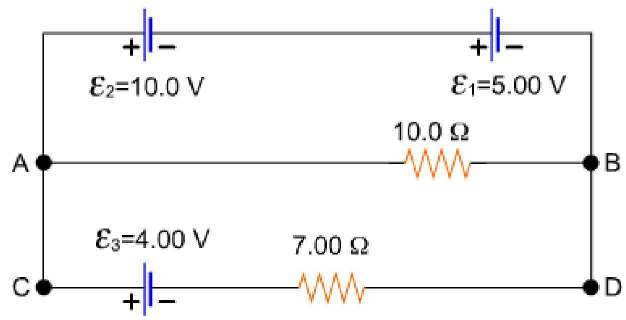
19.1 (a) What is the equivalent resistance of the resistors in the diagram? (b) What amount of current flows through the 30.0 Ω resistor?

(a) __________ Ω
(b) __________ A



19.2 (a) What is the magnitude of the current that flows through branch AB? (b) In what direction does the conventional current flow in branch AB? (c) What is the magnitude of the current that flows through branch CD? (d) In what direction does the conventional current flow in branch CD?

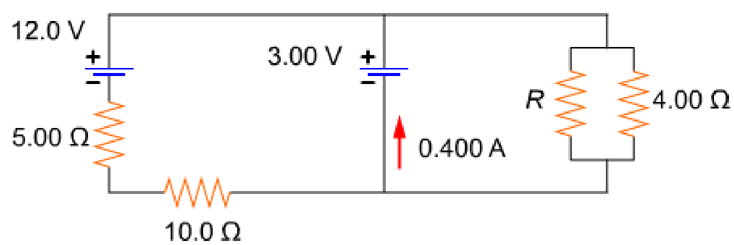
(a) __________ A
(b) ○ To the right ○ To the left
(c) __________ A
(d) ○ To the right ○ To the left



Section 20 - Kirchhoff's junction rule

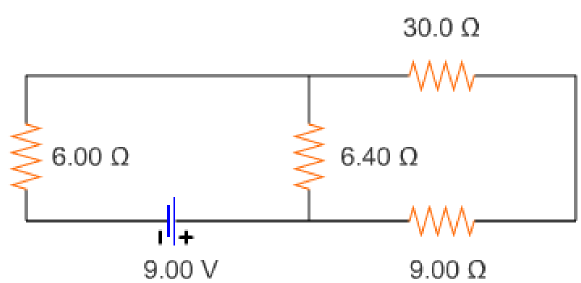
20.1 (a) What is the current through the 10.0 Ω resistor? (b) What is the unknown resistance *R*?

(a) __________ A
(b) __________ Ω



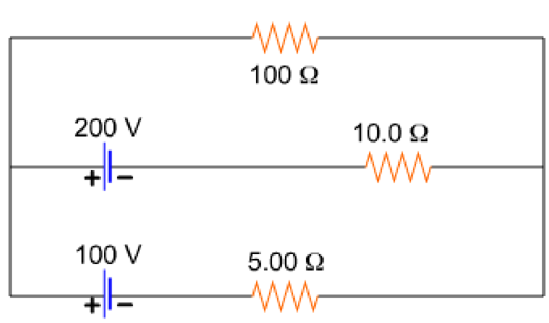
20.2 How much power is dissipated by the 9.00 Ω resistor?

__________ W



20.3 What current flows through the 5.00 Ω resistor? Express your answer to the nearest 0.01 amperes.

__________ A



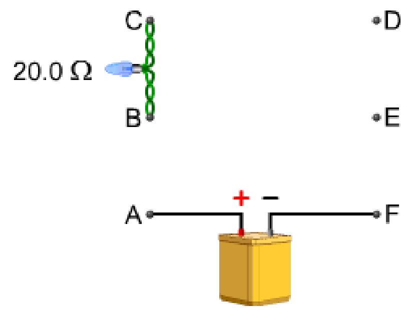
Section 23 - Interactive problem: a complex circuit

23.1 Use the simulation in the interactive problem in this section to calculate (a) the unknown resistance and (b) the unknown emf.

(a) __________ Ω
(b) __________ V

Section 24 - Interactive problem: stringing holiday lights

24.1 Use the information given in the interactive problem in this section to answer the following question. If there is a 20.0 ohm bulb in segment B-C, which component should be placed in the section (a) A-B, (b) C-D, (c) D-E, (d) E-F, (e) B-E in order for all four lights to burn equally brightly.



(a) i. 5.00 ohm bulb
ii. 20.0 ohm bulb
iii. Zero-resistance wire
iv. Nothing

(b) i. 5.00 ohm bulb
ii. 20.0 ohm bulb
iii. Zero-resistance wire
iv. Nothing

(c) i. 5.00 ohm bulb
ii. 20.0 ohm bulb
iii. Zero-resistance wire
iv. Nothing

(d) i. 5.00 ohm bulb
ii. 20.0 ohm bulb
iii. Zero-resistance wire
iv. Nothing

(e) i. 5.00 ohm bulb
ii. 20.0 ohm bulb
iii. Zero-resistance wire
iv. Nothing

Section 26 - Capacitors in series

26.1 There are two capacitors in series, each of 6.5e−5 F. What is their equivalent capacitance?

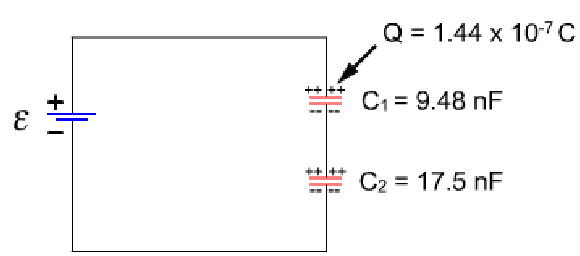
__________ F

26.2 A trio of capactitors, of capacitance 10.0, 20.0 and 30.0 picofarads, are wired in series. What is their equivalent capacitance?

__________ F

26.3 What is the emf of the battery?

__________ V



26.4 You have 12 V battery and a 0.00016 F capacitor in series. You want the potential difference across the capacitor to be 9.0 V. What capacitance should be placed in series with this capacitor to obtain the desired potential difference across it?

__________ F

26.5 Capacitors of capacitance 7.00 microfarads and 13.0 microfarads are connected in series with a 12.0 V battery. (a) What is the equivalent capacitance? (b) What is the charge on the 7.00 microfarad capacitor? (c) What is the potential difference across the 13.0 microfarad capacitor?

(a) __________ F
(b) __________ C
(c) __________ V

Section 27 - Interactive problem: capacitors in a camera flash

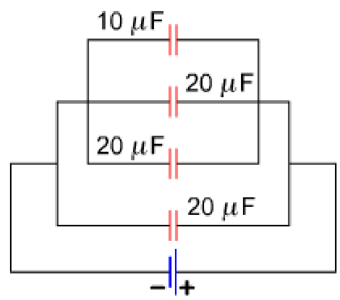
27.1 Use the simulation in the interactive problem in this section to calculate the correct capacitance to illuminate the flash bulb.

__________ F

Section 28 - Capacitors in parallel

28.1 What is equivalent capacitance of the circuit?

__________ μ F



28.2 Three 0.00060 F capacitors are in parallel. What is their equivalent capacitance?

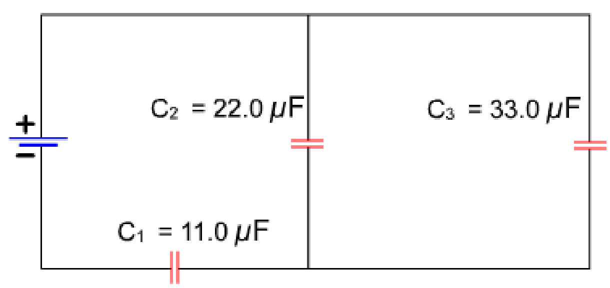
__________ F

28.3 Capacitors of capacitance 7.00 microfarads and 13.0 microfarads are connected in parallel to each other and parallel to a 12.0 V battery. (a) What is their equivalent capacitance? (b) What is the charge on the 7.00 microfarad capacitor? (c) What is the potential difference across the 13.0 microfarad capacitor?

(a) __________ μ F
(b) __________ C
(c) __________ V

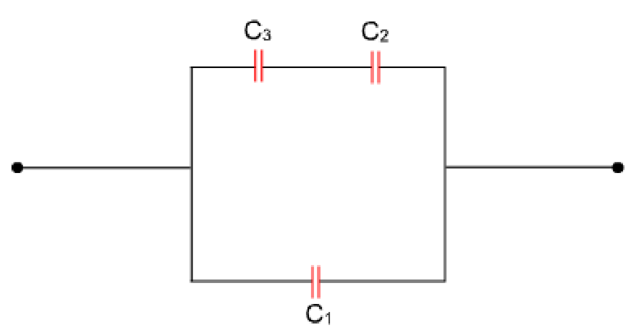
28.4 What is the equivalent capacitance of this circuit?

__________ μ F



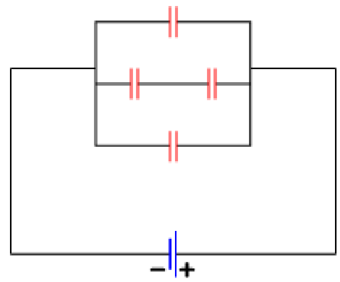
28.5 C_2 has twice the capacitance of C_1, and C_3 has three times the capacitance of C_1. What is the equivalent capacitance?

- ○ $(4/3)C_1$
- ○ $(3/2)C_1$
- ○ $(11/5)C_1$
- ○ $6C_1$



28.6 Each of the capacitors in the circuit has a capacitance of *C*. What is the equivalent capacitance of the circuit in terms of *C*?

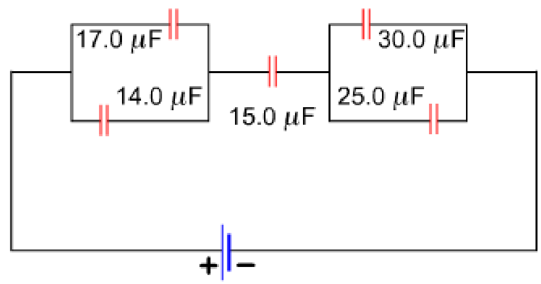
○ 3*C* ○ 5*C* ○ 2.5*C*



28.7 The equivalent capacitance of *n* identical capacitors connected in parallel is 121 times greater than the equivalent capacitance of the same capacitors in series. How many capacitors are being connected?

____

28.8 Any of the capacitors in the diagram will break down (become a conductor) if the potential difference across it exceeds 100 volts. (a) Which capacitor breaks down first as you increase the emf of the battery from zero? (b) What will the charge be on the capacitor when it breaks down? (c) What is the minimum battery emf at which the first capacitor will break down?



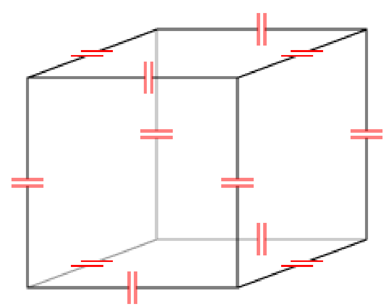
(a) ○ 14.0 μF capacitor
○ 15.0 μF capacitor
○ 17.0 μF capacitor
○ 25.0 μF capacitor
○ 30.0 μF capacitor

(b) __________ C

(c) __________ V

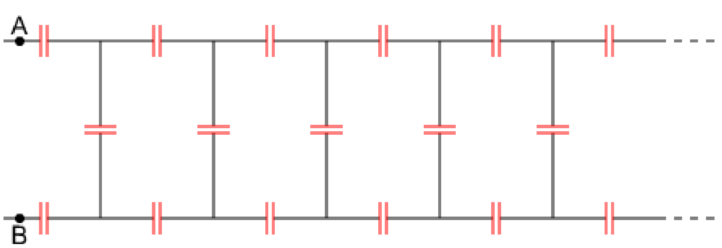
28.9 Twelve identical 5.00 μF capacitors are joined and used to form the edges of a cube. What is the equivalent capacitance between diagonally opposite vertices?

__________ F



28.10 An infinite ladder of capacitors is shown in the diagram. All of the capacitors are identical, and have capacitance of C_0. Determine the equivalent capacitance C_{equiv} between points A and B, as a decimal multiple of C_0. Express your answer in terms of C_0.

__________ C^0



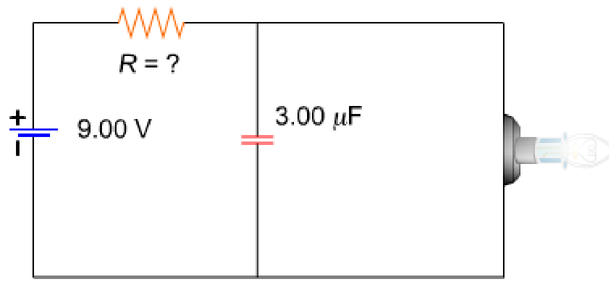
Section 31 - Circuits with resistors and capacitors (RC circuits)

31.1 A camera's flash is powered by a 5.00×10^{-4} F capacitor. When the capacitor is fully discharged, It takes 2.00 seconds to get 99.0% recharged. A resistor is in series with the capacitor. What must its resistance be?

__________ Ω

31.2 Show that the energy in a capacitor in a series *RC* circuit equals $Q\mathcal{E}/2$ when the capacitor is fully charged.

31.3 The circuit diagram for a simple flashing light is shown. The fluorescent light will illuminate when the potential difference across it reaches its breakdown value of 6.00 V. When this value is reached, a current runs through the light and the capacitor discharges completely. Other than flashing when the potential difference across it reaches this value, the light plays no role in the circuit. You wish the light to flash 10.0 times per second. What should *R* be? Assume that the time it takes to discharge the capacitor into the lamp is negligible.



__________ Ω

31.4 You charge an *RC* circuit where the capacitance is 0.50 F, the resistance is 2.0 Ω, and the battery's emf is 25 V. What is the potential difference across the **resistor**, 1.2 seconds after you connect the battery?

__________ V

31.5 A capacitor is connected in series with a 275 ohm resistor. When a switch is closed on a 1.50 V battery connected in series with these components, the capacitor's potential difference increases to 0.875 V in 1.25 microseconds. What is the capacitor's capacitance?

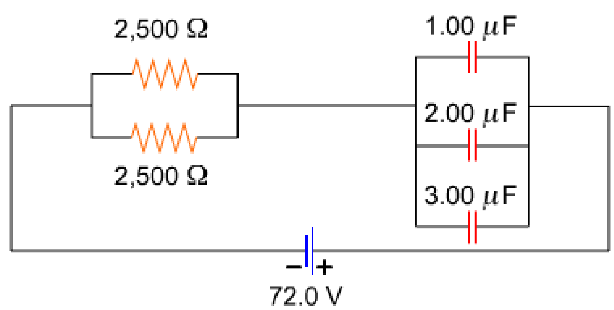
__________ F

31.6 Have you ever wondered how a toaster knows how long to wait before popping the toast up? When you press the toast lever down, an *RC* circuit is connected to a voltage source. Current runs through this *RC* circuit until the capacitor reaches 90 percent of its full charge, at which point another component allows the toast to pop back up. The "darkness" knob on the front of a toaster is really a *variable resistor*, a type of resistor whose resistance changes between a minimum and maximum setting as you turn the knob. The *RC* circuit is made up of this variable resistor and a regular 1.1e−5 F capacitor. What is the (a) maximum and (b) minimum resistance that the variable resistor needs to provide in order for the "toasting" to last between 30 and 90 seconds?

(a) __________ Ω
(b) __________ Ω

31.7 In equilibrium, what is the charge on the (a) 1.00 μF capacitor, (b) 2.00 μF capacitor, and (b) 3.00 μF capacitor?

(a) __________ C
(b) __________ C
(c) __________ C



Section 32 - Discharging a capacitor in an RC circuit

32.1 Before adjusting an electronics experiment, you want to make sure a capacitor you've been using is fully discharged. (Otherwise, you might get a nasty shock!) During the experiment, the 3.5e−5 F capacitor was charged by a 5.0 V potential difference. When you turn the experiment off, the capacitor discharges through a 2.0×10^4 Ω resistor. Mathematically speaking, the capacitor never **fully** discharges; however, practically speaking, we might consider the capacitor to be "fully discharged" when the charge left on it is less than the magnitude of the charge of a single electron: -1.6×10^{-19} C. How long do you have to wait after turning the experiment off for this to occur?

__________ s

32.2 A capacitor of capacitance *C* has an initial charge *Q* and is connected in series with a resistor of resistance *R*. In terms of these variables, how long does it take for the capacitor to lose half of its initial charge?

- ○ *RC*ln(1/2)
- ○ ln(1/2)/*RC*
- ○ ln(2)/*RC*
- ○ *RC*ln(2)

32.3 A 15.0 pF capacitor with an initial charge of 17.0 μC is in series with a 995 Ω resistor when a switch in the circuit is closed. (a) What is the maximum current that flows through the resistor? (b) 25.0 nanoseconds after the switch is closed, what is the current flowing through the resistor? (c) How much charge is on the capacitor at that time?

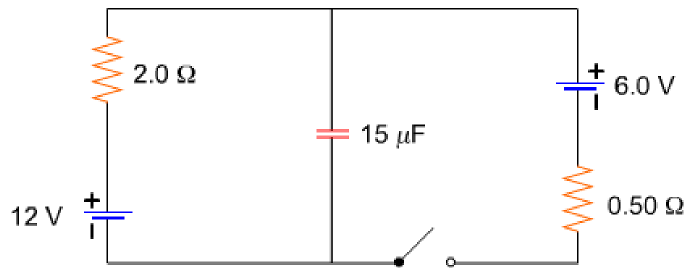
(a) ________ A
(b) ________ A
(c) ________ C

32.4 How many time constants have elapsed at the moment that a discharging *RC* series capacitor has only 0.500% of its initial charge? (Hint: The answer is not a whole number.)

____

32.5 After being left open for an hour, the switch in the circuit shown is closed. 15 minutes after it is closed, (a) what current, if any, is flowing through the 2.0 ohm resistor? (b) What is the potential difference across the capacitor at this time?

(a) ________ A
(b) ________ V

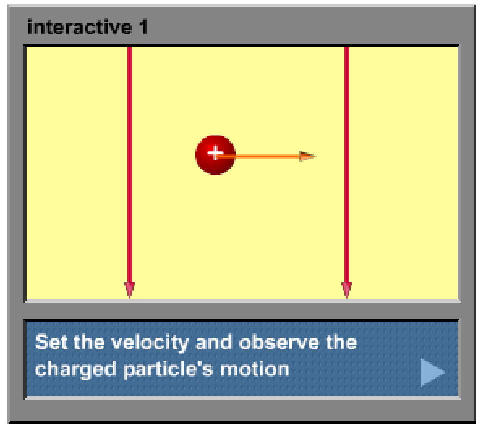


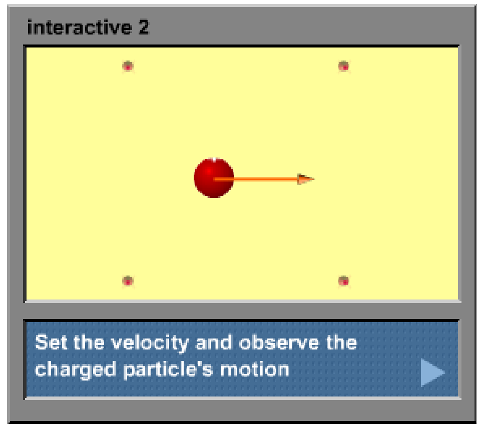
30.0 - Introduction

Humankind has long been familiar with magnets, objects possessing “north” and “south” poles that can attract or repel certain other objects. The ancient Greeks understood their properties, and the word “magnet” itself likely originates from the Greek region of Magnesia, where naturally occurring magnets are found. Early navigators learned to steer their ships with the aid of magnetic devices that were the forerunners of today’s compasses. The natural world also takes advantage of magnetism. A notable example of this is *Aquaspirillum magnetotacticum*, a bacterium that synthesizes tiny magnets that help it determine which way to move. We are a little worried about these creatures: Since the Earth’s magnetic field changes its orientation every several hundred thousand years or so, the microorganisms may find themselves unintentionally heading away from dinner one day.

Much more complex creatures, namely scientists, employ magnets as well. In the 16th and 17th centuries, they learned to create their own magnets and used them to study the Earth’s magnetic field. In the 19th century, scientists began the crucial work of piecing together a more complete picture of the relationship between magnetic fields and electric currents.

Despite centuries of practical use, magnets and their fields still pose mysteries. For instance, scientists cannot definitively establish the cause of the Earth’s magnetic field, and they are only able to speculate about why its direction periodically changes. In addition, physicists continue their quest for the magnetic monopole: a magnet with just one pole. As you proceed in your studies in this chapter, remember that you are in the good company of other fine minds who have found the workings of magnets to be an area of continuing fascination.





You can begin your exploration of magnetic fields by launching the simulations to the right that show the effect of a magnetic field on the motion of a charged particle. These two simulations are the same except for the initial viewing angle, the angle at which you view what is occurring in the simulation. In both, you control the initial velocity of a positively charged particle that moves in a magnetic field, represented by magnetic field lines. In the illustration for Interactive 1, you see that the magnetic field points straight down the screen and the initial velocity vector points to the right, perpendicular to the magnetic field. As the particle moves you will see, represented as a purple vector, the force exerted on it by the magnetic field.

Clicking on Interactive 2 launches the same simulation but with the viewing angle rotated $90°$. Here, the magnetic field points directly toward you, and you are seeing the heads of the field lines. With either simulation, you can change the viewing angle by using the slider provided, and see either of these points of view. In the simulations you will also see a magnetic field meter that is there principally to help you understand the changing perspective as you manipulate the viewing angle slider.

Launch the upper simulation and conduct some experiments. Does the moving particle travel along a straight line or a curve? To answer this question, you will need to change the viewing angle, in the process seeing why a three-dimensional view of a charged particle moving through a magnetic field is so useful.

As you study the path, answer two more questions. First, is the particle’s speed changing? Second, is it accelerating? For the second question, recall that acceleration measures the change in velocity, which is a vector.

You can also consider the relationship of the directions of the various vectors you see. What is the relationship between the force and velocity vectors? Are they parallel or perpendicular? What is the angle between the force vector and the magnetic field? Again, the viewing angle slider proves a useful tool. You can only change the direction of the initial velocity when the viewing angle is set to the far right so that the magnetic field is pointing straight down the screen; this makes it easy to see the angle between the velocity vector and the magnetic field vector.

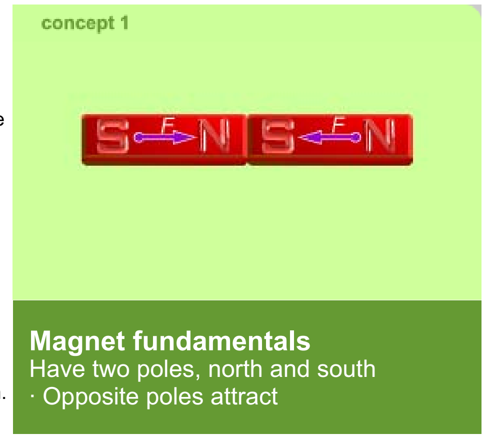
As another experiment, set the velocity to zero. Does the magnetic field exert a force on the particle? If the magnetic field exerts a force, the particle will accelerate. How does this compare to what would occur if the stationary charged particle were in an electric field? (Note: We ignore other forces, notably gravity and air resistance, in these simulations.)

We have asked a lot of questions above. Answer as many as you can now, and prepare to explore magnetism in depth in this chapter.

Magnet: An object that creates a magnetic field and exerts a magnetic force. All known magnets have two poles.

This boy is holding a U magnet. The magnet suspends a few paperclips.

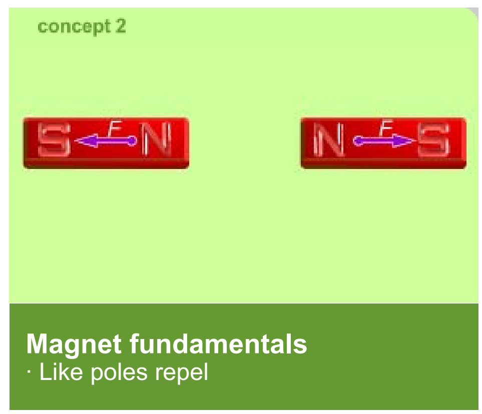
Just like electrical charges, magnets create fields and exert forces. All magnets are dipoles, meaning they have two poles: a north pole and a south pole. As with electrical charges, magnetic opposites attract and likes repel; a north pole attracts a south pole, and a pair of like poles, such as two south poles, repel each other.



A rectangular magnet like the one shown to the right is called a bar magnet. It has a magnetic pole at each end. If you were to bend this magnet into a “U,” you would create a horseshoe magnet like the one the boy is holding in the photo above. Even though you have changed the magnet’s shape, the poles remain at the two extremities.

The attraction and repulsion of magnetic poles can be demonstrated with a pair of bar magnets. If you have two such magnets, position them so that their opposite poles are adjacent. What happens? They attract each other. You can see this illustrated in Concept 1. Now place them so their like poles are next to each other; the magnets will repel each other. This is shown in Concept 2.

The strongest and most frequently found form of magnetism is called *ferromagnetism*. Only certain types of material (iron is a notable example) exhibit this form of magnetism. Everyday videotapes reveal a common application of ferromagnetism: A movie is encoded as a pattern of small magnets on the videotape, and the VCR “reads” this data. Computer hard drives function in a similar way.

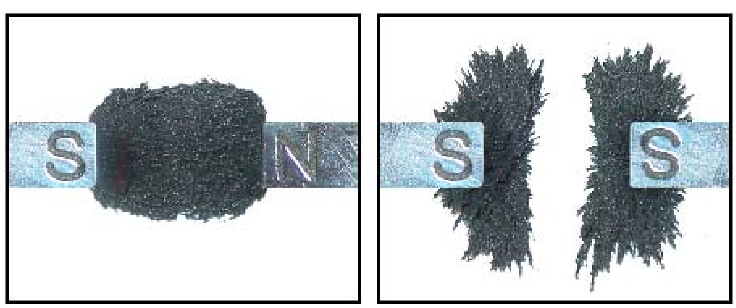


You may have noticed that magnets can “stick” to surfaces that do not initially exhibit magnetism. For instance, the exterior of your refrigerator is typically not magnetic, and yet if it is made of a ferromagnetic metal, magnets will stick to it. A magnet sticks because it is able to induce a temporary magnetic field in the refrigerator’s surface. The permanent and temporary magnets then attract each other.

Another way to observe this phenomenon is to attach a metal paper clip to a magnet and then attach a second clip to the first clip. The two clips will attract each other. When you remove the original magnet, however, the clips will no longer attract one another because the fields of the paper clips disappear when they are removed from the influence of the magnet’s field.

No one has created a magnet with just one pole (a monopole). Magnetic poles always come in pairs. If you take a bar magnet and cut it in half, you will create two magnets, each with a south and a north pole. Cut each of those in half and you will have four magnets, each with two poles. In principle, if you could cut the magnet into its constituent atoms, each atom would have its own magnetic field.

30.2 - Magnetic fields



Iron filings line up with the magnetic field.

Two magnets can attract or repel each other without touching: They exert a force at a distance. Magnetic fields surround magnets. Like an electric field, the magnetic field is a vector field. It has a strength and a direction at every point. The letter $\mathbf{B}$ represents the magnetic field and the unit for magnetic field strength is the *tesla* (T).

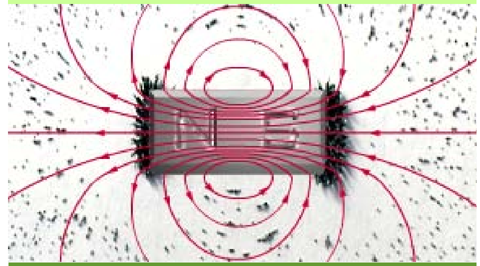
The photographs above show the alignment of iron filings gathered around the poles of two pairs of magnets held close together. The filings in both photos align with the magnetic fields between the poles. The photograph on the upper left shows the filings between a south pole and a north pole, which attract each other. The filings are connected and aligned in the same direction as the magnets because the magnetic field lines pass directly from the north pole to the south pole. In the photograph on the upper right, two south poles are repelling each other, which causes the filings to separate.

Like electric fields, magnetic fields can be diagrammed with field lines. We have superimposed a drawing of magnetic field lines on the illustrations to the right. You can see how the orientations of bits of iron filings reflect the direction of the field at a number of places around the magnet.

As the diagram shows, the field lines run outside the magnet from the north to the south pole, and they always form closed loops. As with an electric field, the closer the lines, the stronger the field. The strength of the field decreases with distance from the magnet.

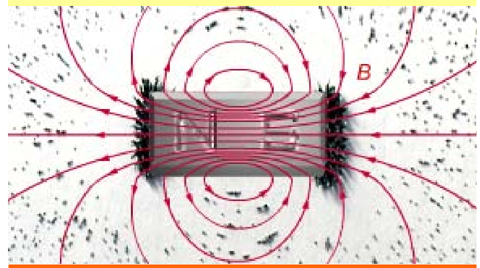
The table in Equation 2 shows approximate magnetic field strengths near some magnetic objects. The magnetic fields you encounter every day are much less than one tesla. The most powerful lab magnets have field strengths of tens of teslas. However, human technology cannot create magnetic fields anywhere near the field strengths exhibited by a class of highly magnetized neutron stars called *magnetars*.

concept 1



Magnetic field
Region where magnet exerts force
External lines point from north to south

equation 1



Magnetic field
B is symbol for magnetic field
Units: teslas (T)

equation 2

	Approximate magnetic field (T)
1 foot from hairdryer	10^{-7}
Under power distribution line	10^{-5}
Earth's surface	10^{-4}
Refrigerator magnet	10^{-2}
Medical MRI	1.5
Most powerful lab magnets	60-100
White dwarf star surface	100
Magnetar star surface	10^{11}

Magnetic field strength near magnetic objects

30.3 - Physics at work: lodestones

Lodestone: A naturally occurring magnetic rock found in various regions of the world.

The Chinese first used lodestone magnets in their navigational compasses in the twelfth century. Later explorers, including Columbus, carried lodestones to “recharge” their iron-based compasses. To do so, they would pass the needle of a compass near the lodestone to strengthen its magnetic properties.

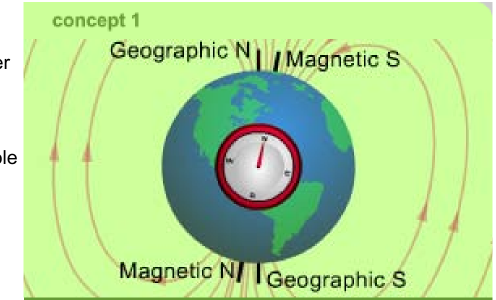
A lodestone consists of two materials, one with a high *saturation magnitude*. Saturation magnitude is the measure of the maximum strength of the magnetic field that can be induced in a substance. Iron, for example, has a high saturation magnitude. The other material has high *coercivity*, which is the measure of how well a substance retains its magnetic field. Pure iron loses its magnetic properties relatively easily, while a substance such as magnetite does not.

The combination of these two materials results in a strong and durable magnetic source. One substance supplies the strength, the other supplies durability.

concept 1

Lodestones
Natural magnetic rock combines materials with
· Strong magnetic field
· Long-lasting magnetic field

30.4 - The Earth and magnetic fields



The Earth and magnetic fields
Earth is huge magnet
· North Pole ≈ south pole of magnet
· South Pole ≈ north pole of magnet

The Earth acts like a huge magnet with an accompanying magnetic field. Compasses take advantage of this field: Their needles are mounted so that they can rotate freely and align with it. Once a compass has established the north and south directions, other directions can be established relative to them.

The Earth's **magnetic** south pole is located near its **geographic** North Pole. This means that the north pole of the magnet in a compass points north, since the south pole of the planetary magnet attracts it. The geographic poles lie on the planet's axis of rotation.

Scientists still do not know exactly how our planet produces its magnetic field, though the prevalent theory is that Earth's magnetism arises from processes that occur in its core. Scientists have arrived at this conclusion because of the fundamental differences between the cores of the Earth and the Moon. The Earth has a molten metallic core and a magnetic field, while the Moon has a solid core and no magnetic field. Many scientists believe the large-scale liquid flow that occurs inside the Earth's core causes its magnetic field.

While scientists are still puzzled by the "why" of Earth's magnetic field, they do know that the locations of its magnetic north and south poles change. The magnetic south pole is currently about 1900 km away from the geographic North Pole of the Earth. The magnetic north pole resides in the ocean south of Australia, not even in the continent of Antarctica! These positions are not fixed; shifts of several degrees in their locations have been measured over the last century. The poles do more than wander, they also reverse their orientation. The last switch occurred a few hundred thousand years ago, when magnetic north became magnetic south, and vice versa.

Scientists deduce this change in orientation by analyzing parts of the ocean floor that originate as magma (molten rock) emerging from cracks in the Earth's crust. Over a period of millions of years, vast quantities of solidified magma have been deposited on the ocean floor. As magma congeals at any point in time, the iron it contains "records" the orientation of the Earth's magnetic field. New magma flows force earlier deposits apart, which separates them and enables scientists to establish their sequence. By examining different sections of the rock, scientists can deduce the direction of the magnetic field at various times throughout Earth's history.

The current strength of Earth's magnetic field is about 5×10^{-5} teslas, and that strength is decreasing at about 0.07 percent per year. If it continues to weaken at this rate, it will be reduced to only 1 percent of its present value in 6,500 years. Since the magnetic field helps to shield our biosphere from cosmic rays and charged particles from the Sun, this could be a matter of concern.

Scientists theorize that the poles "flip" alignment after the field passes through a state with zero magnitude. Because there are no signs of massive mutations in the fossil record that date from the period of the last "flip," perhaps the results are not as severe as one might fear. No one knows exactly how long the zero-field condition exists, whether five years, 50 years, or 1000 years. As the NASA website above has stated, "Stay tuned…"

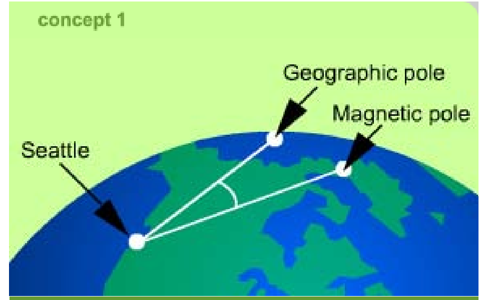
30.5 - Physics at work: compasses and the Earth

A magnetized compass needle aligns with the Earth's magnetic north and south poles. The needle's north pole, sometimes distinguished by a bright color as in the photograph above, points toward the Earth's magnetic south pole. However, the Earth's magnetic poles are not at the same locations as its geographic poles. For instance, in 1831, British explorer James Clark Ross located the magnetic south pole off the coast of Canada (remember, the magnetic south pole corresponds to the geographic North Pole). He tried to duplicate this feat for the magnetic north pole near Antarctica, but failed due to weather and ice. Nonetheless, many regions of that continent (such as the Ross Sea) are named after him.

In mountains near Seattle, Washington, the magnetic declination is about 20°.

The angle between the directions to the magnetic and geographic poles is called the magnetic *declination*. Many topographic maps provide the local angle of declination so that hikers and others can compensate for it and orient their compasses to geographic north. The first sketch on the right shows the declination for Seattle, Washington: It is approximately $20°$.

Scientists and navigators also found that compasses could be used to estimate latitude (north/south position on the planet). Although today's compasses are designed to move in a horizontal plane, the needles in some early instruments were allowed to rotate freely in all directions (like a bar magnet suspended by a string tied around its middle). When carried to a magnetic pole of the Earth, the magnetic needles of these compasses would point straight up or down; conversely, when near the equator, they would point almost horizontally. By measuring the angle of the needle from the horizontal, navigators could estimate their latitude.

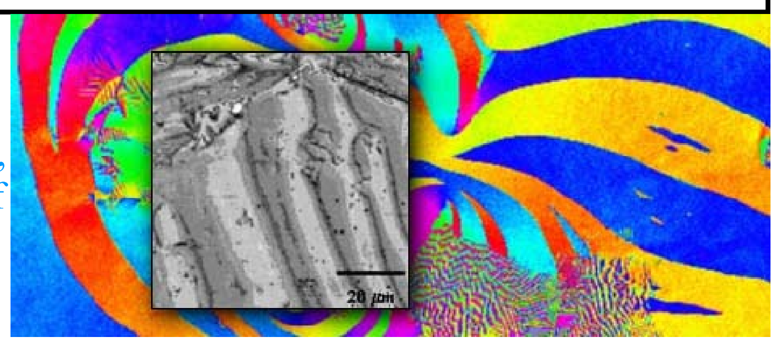


Declination
Angle between magnetic, geographic pole

Latitude
Can be estimated with a dip needle

30.6 - Ferromagnetism

Ferromagnetism: A strong magnetic effect exhibited by the atoms of certain elements, notably iron. It is the cause of the magnetic field of commonly used magnets.

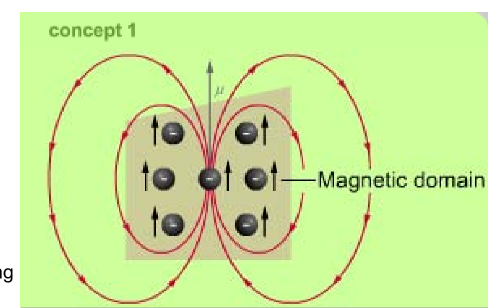


Color image: magnetic domains in an amorphous ferromagnet. Inset: magnetic domains in heat-treated carbon steel.

Ferromagnetism is the basis of the most familiar type of magnetic devices. When you see a magnet affixed to a refrigerator door, you are witnessing the results of ferromagnetism. "Ferro" comes from the Latin word for iron, since iron is the most common element to exhibit this property. Ferromagnetism is present in the magnets that can be purchased in toy or hardware stores.

Nickel and cobalt are two other common elements with atoms that exhibit ferromagnetism, though to a lesser extent. Gadolinium and dysprosium are two exotic elements that can be used to make strong magnets that perform well at low temperatures.

In order to explain the source of ferromagnetism, we have to pay a quick visit to a concept developed in quantum theory. The classic model of atoms − electrons revolving about a central nucleus − does not suffice. Ferromagnetism is caused by a property of electrons called *spin*. For the purposes of this discussion, we say that spin is used to describe the angular momentum of an electron due its rotation, akin to the angular momentum of the Earth due to its rotation about its axis. (This is only an analogy. A more detailed, much less mechanistic description requires the context of quantum theory.) This is one form of angular momentum possessed by the electrons.



Inside a domain
Electron spins align due to exchange coupling
· Result is magnetic domains

They also possess *orbital angular momentum* due to their motion about the nucleus. Both contribute to the magnetic moment of the electron, but only spin is relevant to the discussion of ferromagnetism. The magnetic moment vector helps quantify how much torque a magnetic dipole will experience in a magnetic field; the larger the moment, the greater the torque. The moment points from the south pole to the north pole of the dipole.

Spin means that each electron has its own magnetic field. It can be considered as acting like a bar magnet, a dipole with north and south poles.

In many materials, the spin of each electron cancels out that of another electron in the same atom with which it is “paired,” resulting in no net magnetic field. Ferromagnetism results from the net magnetic field created by unpaired electron spins. For example, each iron atom contains four electrons with uncancelled spins, giving the iron atom a net dipole moment.

In ferromagnetic materials, the spins of the electrons of one atom interact with the spins of electrons of neighboring atoms. This interaction, called *exchange coupling,* is quantum mechanical in nature. As a result of exchange coupling, the magnetic dipoles of atoms within a ferromagnetic material tend to align in the same direction.

When the magnetic dipole moments of ferromagnetic atoms align, the result is a *magnetic domain*, a region in a material where there is a net magnetic field. Magnetic domains can be seen under powerful microscopes, often with the aid of a technique called magnetic force microscopy. In the color micrograph above, domains are given false colors to be more easily seen in the image. Typical domains have a diameter about one-third that of a human hair. This can be best seen in the small inset micrograph that shows the magnetic domains in a sample of carbon steel. Even small amounts of magnetic materials contain vast numbers of domains. A domain may contain 10^{12} to 10^{15} atoms, but since a cubic centimeter of iron contains about 2.5×10^{19} atoms, there are still a large number of domains.

There is a good question here: Why are not all iron objects magnets? The reason is that although domains have magnetic fields, they point in random directions unless they have been subjected to an external magnetic field. This is shown in Concept 2. The random nature of their directions means there is no net magnetic field in the substance.

However, when the ferromagnetic material is placed in an external magnetic field, a process occurs that results in the material gaining its own overall magnetic field. This happens because the external field exerts a torque on the magnetic dipoles.

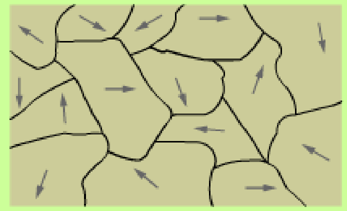
The magnetic field overcomes the tendency of the dipoles to stay aligned with their neighbors, and causes them to align with the field. Unaligned dipoles bordering domains that are aligned with the field steadily become “converted” to the new orientation. This causes the domains aligned with the external field to expand. This is shown in Concept 3, where the domains shown in Concept 2 have combined to form three domains.

Since there is a tendency for the dipoles to remain aligned, when the external magnetic field is removed, the domains (and the dipoles that make them up) remain aligned. This means the material now has its own magnetic field. In short, a magnet has been created.

An external magnetic field can cause non-ferromagnetic materials to develop their own magnetic field. These other forms of magnetism are called diamagnetism and paramagnetism. However, ferromagnets are distinct and quite useful for two reasons. First, a ferromagnet retains its magnetic field after the external field is removed. A refrigerator magnet, or the information stored on a VCR tape, both rely on the longevity of ferromagnets. The other forms of magnetism disappear after the external magnetic field is removed.

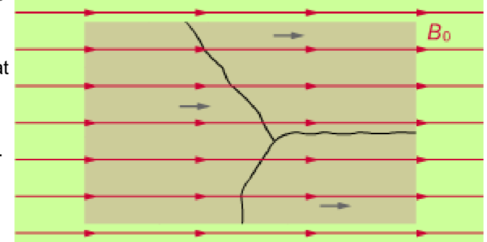
Second, the magnetic field created by ferromagnetic materials is typically orders of magnitude stronger than that found in the other forms of magnetism in most materials. This makes ferromagnets useful for a wide range of everyday applications.

concept 2



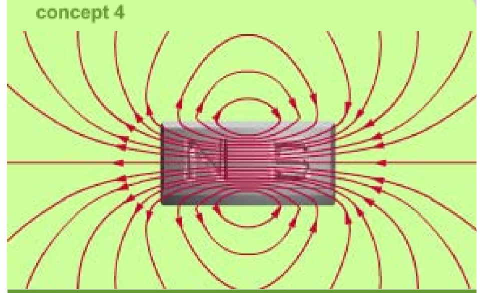
Random domains
Unaligned domains mean no net magnetic field

concept 3



In external magnetic field
In external magnetic field, B_0
· domains aligned to external field grow

concept 4



Magnetic field removed
Substance maintains magnetic field

30.7 - Magnetic fields and charged particles

Magnetic fields exert forces on **moving**, electrically charged particles.

This phenomenon makes for a good demonstration in a physics class. A current-carrying wire is placed near a magnet. The magnet exerts a force on the electrons moving in the wire, which causes the wire to move toward or away from the magnet. When the current is turned off, the magnetic field stops exerting a force on the wire.

TV tube. A gun (far left) accelerates electrons across a potential difference. Electromagnets (brown) steer the moving charges to light up screen pixels.

The images on the screens of traditional televisions and computers are the result of electrons being accelerated by an electric field, subsequently being “steered” by magnetic fields, and then striking the screen to create light of different colors at specific locations. If you were to place a magnet close to such a system, you would distort the image. However, we do not recommend you do this, as it could cause expensive or irreversible damage!

Four factors determine the amount of force exerted by a magnetic field on a moving particle. They are the particle's charge and speed, the strength of the magnetic field and the angle of intersection between the particle's velocity and the magnetic field. The force is greatest when these two vectors are perpendicular, and zero when they are parallel.

When a charged particle is surrounded by an external **electric** field, the electric force on the charge is exerted along the field lines. The electric field exerts a force on the charge whether it is moving or stationary. In contrast, **magnetic** fields only exert a force on **moving** charged particles, pushing them neither in the direction of the particle's motion nor along the lines of the field. The force exerted on a moving charge by a magnetic field is perpendicular to **both** the particle's velocity and the direction of the field.

You see this illustrated to the right, and you also experienced it when you used the two simulations in the introduction to this chapter. In Concept 1, the same phenomenon is shown from two different vantage points.

In both illustrations, a positive charge is moving through a magnetic field. In the view labeled "side view," the magnetic field points away from you. This is depicted with ×'s, which represent the field lines viewed from behind. This view is used to best show the direction of the force: It is perpendicular to both the field and the velocity vectors.

In the view labeled "front view," your viewpoint has been rotated $90°$ so the field appears parallel to the screen. The force vector in the front view points toward you and is represented by a dot. You are looking at "the business end" of the vector's arrow.

A right-hand rule is used to determine the direction of the force. The front view makes it easier to see, so we use that in Concept 2. To apply the right-hand rule, the fingers of a flat hand start out pointing in the same direction as the velocity vector of the charge. Then they curl, so that the fingers point in the same direction as the magnetic field. The fingers wrap **from** the velocity vector **to** the magnetic field vector.

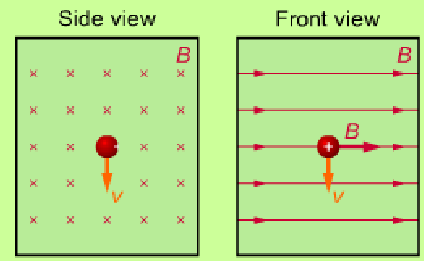
For a positive charge, the thumb points in the direction of the force exerted on the moving charge. The thumb points **opposite** to the direction of the force for a **negative charge**. In other words, with a negative charge, you apply the same rule, but reverse the results.

The equation to determine the force is shown to the right. The force equals the charge times the cross product of the velocity and magnetic field vectors. The cross product is calculated with the sine, as shown to the right.

To determine the amount of force, multiply the absolute value of the charge (its positive value), the charge's speed, the field strength, and the sine of the angle between the velocity and field vectors. This angle is shown in the illustration for the equation. When calculating the force magnitude, you use the smaller, positive angle between the velocity and magnetic field vectors. For instance, in the Equation 1 diagram the angle is $90°$, not $-90°$ or $270°$.

As mentioned earlier, the unit for magnetic field strength is the tesla. One tesla equals one newton·second per coulomb·meter, or N·s/C·m. In other words, one coulomb of charge traveling at one meter per second through a magnetic field having a strength of one tesla experiences one newton of force. A tesla is a rather large unit (remember that one coulomb is a lot of charge), so the smaller unit *gauss* (G) is fairly common. Ten thousand gauss equals one tesla. As mentioned earlier, the Earth's magnetic field is about 5×10^{-5} T, which equals 0.5 G.

concept 1

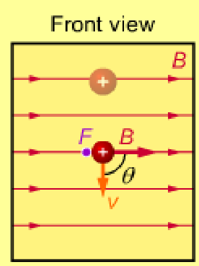


Magnetic field, charged particle
Magnetic field exerts force on moving charge
· Force perpendicular to velocity, field

concept 2

Right-hand rule
Determines direction of force
· Fingers curl from v to B
· Thumb shows force on positive charge

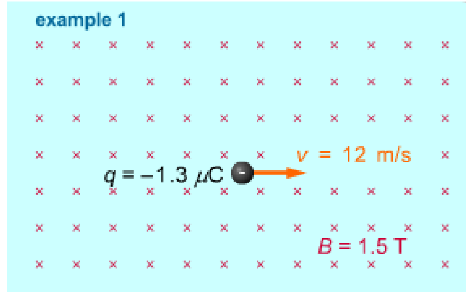
equation 1



Force on charge moving through magnetic field

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}$$
$$F = |q|vB \sin\theta$$

$\mathbf{F}$ = magnetic force on particle
q = charge
$\mathbf{v}$ = velocity
$\mathbf{B}$ = magnetic field
θ = angle between velocity and field



The particle is moving to the right. What are the magnitude and direction of the force on it?

$\theta = 90°$

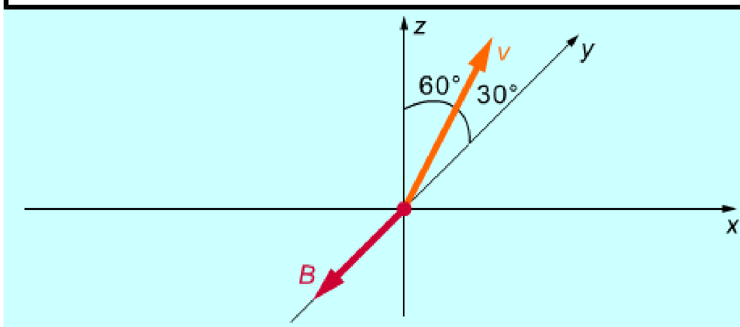
$F = |q|vB \sin\theta$

$F = |-1.3 \times 10^{-6} \mathrm{C}|(12 \frac{\mathrm{m}}{s})(1.5\mathrm{T})(\sin 90°)$

$F = 2.3\times10^{-5}$ N

directed down (opposite to thumb)

30.8 - Interactive checkpoint: a charge in a magnetic field

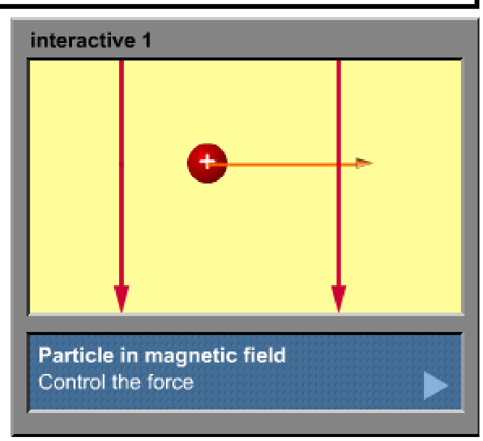


An electron is traveling in the *yz* plane at an angle of 60.0° from the positive *z* axis and 30.0° from the positive *y* axis as shown. Its speed is a constant 365 m/s when a uniform magnetic field of 1.22×10^{-3} T is turned on, pointing in the negative *y* direction. State the acceleration of the electron in the instant after the magnetic field is turned on as a vector in rectangular notation.

Answer:

a = (____ , ____ , ____) m/s^2

30.9 - Interactive problem: charged particle moving in a B field



In this simulation you can change two properties of a particle: its charge and its velocity (both speed and direction). We ask you to alter the velocity of the particle in order to achieve certain forces.

In the initial view of the magnetic field it is directed straight down, and the charge starts near the middle of the screen. You can use the slider in the control panel to change the viewing angle. Press GO to start the simulation, and press RESET whenever you want to change the particle's initial velocity or charge.

You can only change the direction of the initial velocity when the viewing angle is set to the far right and the magnetic field is pointing straight down. This orientation makes the angle between the velocity vector and the magnetic field vector easy to see. You change the direction of the velocity vector by dragging it with the mouse.

Your first task is to set the direction of the particle's initial velocity so that no force is exerted on the charge by the field. No shortcuts! You could solve this by setting the velocity to 0 m/s, but we want you to solve the problem by setting the velocity vector's direction correctly. When the field exerts no force, both the direction and speed of the particle will remain constant.

Note that after you change the charge or the velocity vector (speed or direction), but before you press GO, the force meter tells you what the force on the particle will be, not what it is. The force is zero whenever the particle is not moving, but we thought you would appreciate seeing immediate feedback on your changes, without having to launch the particle every time.

What is the relationship between the directions of the field and the velocity vector when the force equals zero? (Note: There are actually two directions in which you can set the velocity vector to achieve this, one $180°$ opposed to the other.)

For your next challenge, set the parameters in the simulation to maximize the amount of force on the particle. As with your first task, there are two directions that result in a maximum force on the charge. You will find yourself clicking the up and down arrows on parameters like “Charge” until you reach their maximum value in the simulation. The maximum amount of force you can achieve in this simulation is 1.68×10^{-8} newtons.

For your final task, change the charge from positive to negative while keeping the initial velocity the same. What effect does this have on the magnitude and direction of the force?

If you are surprised by the results of any of these configurations, refer back to the section that discussed the effect of a magnetic field on a moving charged particle.

30.10 - Determining the strength of a magnetic field

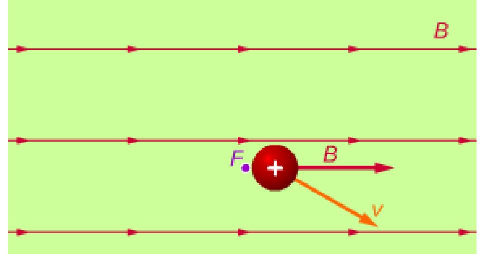
A magnetic field exerts a force on a moving charged particle. When the charge, velocity and field are known, the direction and magnitude of that force can be calculated. This process can also be reversed: Just as an electric field can be measured using a test charge, the strength of a magnetic field can also be determined using a test charge. The process is similar, but with a couple of twists.

In an electric field, a stationary positive test charge is used to determine the electrostatic force at a given point. The amount of force is then divided by the test charge to determine the electric field strength. The direction of the force on the test charge indicates the direction of the field.

In a magnetic field, the force must be determined using a **moving** charge. Physicists fire a charged particle through a magnetic field and measure how the force exerted on it by the field alters its velocity. They then use Newton’s second law, together with the observed acceleration and the mass of the particle, to calculate the force the field exerts on it.

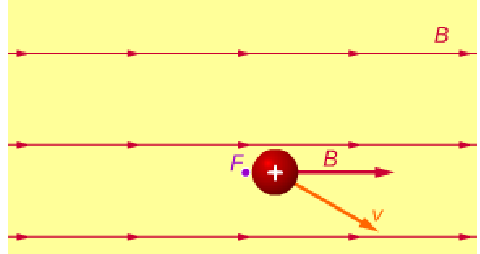
The angle between the velocity and the magnetic field vectors must also be known to determine the field. (A compass can be used to determine the direction of the field.) Once the force and angle are known, they can be used in the equation shown on the right to find the strength of the magnetic field. This is the same equation shown earlier to determine the amount of force exerted by a given magnetic field, but here it is solved for the field strength.

concept 1



Measuring a magnetic field
Use a moving test charge
Measure acceleration of charge
Newton’s 2nd relates acceleration, force
With force known, field can be found

equation 1

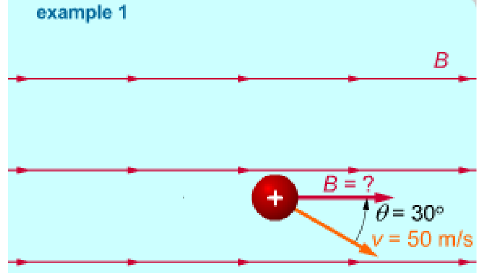


Measuring a magnetic field

$$B = \frac{F}{|q|v\sin\theta}$$

B = magnetic field strength
F = force, q = charge
v = speed
θ = angle between velocity and field

example 1



The force is 4.0×10^{-5} N and the charge q is 2.0 μC. What is B?

$$B = \frac{F}{|q|v\sin\theta}$$

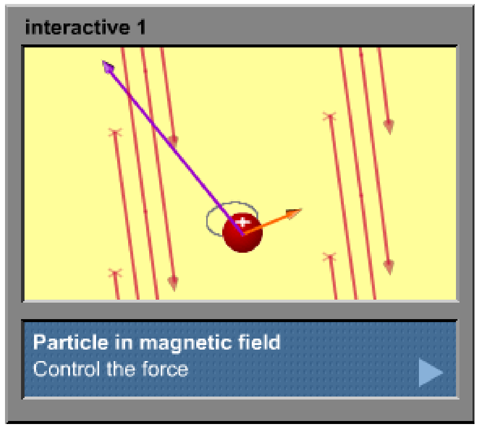
$$B = \frac{4.0\times10^{-5}\,\text{N}}{|2.0\times10^{-6}\,\text{C}|(50\,\frac{\text{m}}{\text{s}})(\sin 30°)}$$

$$B = 0.80\ \text{T}$$

30.11 - Interactive problem: B field strength and particle motion

In this simulation, you want the magnetic field to exert a force of 1.16×10^{-8} newtons on the positively charged particle. Set the particle's speed v (the maximum possible value is 520 m/s) and the magnetic field strength B (the maximum possible value is 3.00 T) to cause the required amount of force, using the controllers provided in the control panel. You will find that you need to adjust both the speed and the field strength upward from their initial settings to achieve this force.

If you have any trouble achieving the desired force, review the discussion in the previous section on determining the strength of a magnetic field. If you have any questions about how to use the simulation, see the previous interactive problems in this chapter for complete instructions.



30.12 - Physics at work: velocity selector

A velocity selector is a device that allows charged particles moving only at a specified speed to pass through it. It uses a combination of an electric field and a magnetic field to "trap" particles moving at other speeds. In addition to being a useful tool, a velocity selector provides a good way to explore the contrasting effects of electric and magnetic fields.

To the right is a conceptual diagram of a velocity selector for charged particles. Perhaps a scientist wants only electrons moving at 50,000 m/s to pass through. The electric field of the velocity selector points down, and will exert an upward force on an electron. (Remember that the force is opposite to the direction of the electric field because electrons are negatively charged.)

A magnetic field is used to counter the effect of the electric field. The magnetic field strength is set so that the field exerts an equal but opposite force on any electron moving at 50,000 m/s. Since the force exerted on the electron by the electric field is upward, the magnetic field force must point down. Electrons with the required speed will travel horizontally and pass through an aperture at the right end of the selector. Slower or faster ones will experience a net force and be forced up or down. (We ignore the force of gravity, whose effect would be minor for a particle moving at this speed.)

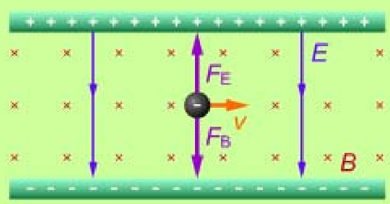
The sum of the forces on the particle is $q\mathbf{E} + (q\mathbf{v}\times\mathbf{B})$. This equation is called the *Lorentz force law*. The strength of the electric force equals $|q|E$. The strength of the magnetic force equals $|q|vB$ (since the velocity is perpendicular to the field). When the forces sum to zero, these two expressions are equal in magnitude. On the right, we solve for the speed and see that it equals the ratio of the electric to the magnetic field strength. As you can see, the charge cancels out.

Can you predict the direction in which this device's magnetic field needs to point using a right-hand rule? The electric force on the electron points up, so we want the magnetic field force to point down. Since the charge is negative, the thumb will point in the opposite direction of the force. This means you want the thumb to point up. The fingers must wrap from the velocity vector to the magnetic field vector. You can conclude that the magnetic field vector must point away from you, which is how it is depicted in the diagram.

In the setup shown here, if a negatively charged electron is moving too fast, the force exerted by the magnetic field will be greater than that exerted by the electric field, and as a result it will pull it down. If the electron is moving too slowly, the electric field will win the contest and pull it up. The example problem to the right shows how to determine the strength of the magnetic field that permits an electron moving at 53,000 m/s to pass undeflected through a uniform electric field with a magnitude of 4.0 N/C.

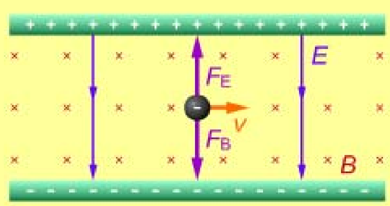
It is worth emphasizing that the name of the device is a *velocity* selector, not a *speed*

concept 1



Velocity selector
Charges of specific speed pass through
Uses electric and magnetic fields
Electric force balances magnetic force

equation 1



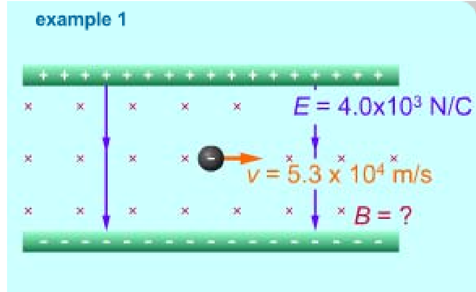
At equilibrium:
$\mathbf{F}_E + \mathbf{F}_B = 0$, so $|q|E - |q|vB = 0$, and

$$v = \frac{E}{B}$$

$\mathbf{F}_E$ = force of electric field
$\mathbf{F}_B$ = force of magnetic field

selector. Direction matters. You may want to consider what magnetic force would be exerted on a charged particle that was moving horizontally, but from right to left as it passed through the selector at the critical speed. Would the magnetic force still cancel out the electric force?

q = charge, v = speed of charge
E = electric field strength
B = magnetic field strength



This electron passes through the velocity selector with its velocity unchanged. What is *B*?

$v = E/B$

$B = E/v$

$B = (4.0\times10^3 \text{ N/C})/(5.3\times10^4 \text{ m/s})$

$B = 0.075$ T

30.13 - Circular motion of particles in magnetic fields

When the velocity of a charged particle is perpendicular to a uniform magnetic field, it causes the particle to move in a circular path. In this section we discuss why this occurs, and we state some properties of that motion.

At the right is a diagram showing a magnetic field (pointing directly into the screen) and a positively charged particle moving in a circular path. Its motion started when the particle was fired into the field along the surface of the page.

Why is the motion circular? First, recall that the magnetic force is always perpendicular to the velocity vector. This means it neither increases nor decreases the speed of the particle. It only changes the direction of its motion.

The force always points toward the center of the circle. You can use the right-hand rule at any point of the circle to confirm this. (The fingers wrap from $\mathbf{v}$ to $\mathbf{B}$, so the thumb points toward the center.)

The magnitude of the force is constant. The quantities that determine the force, q, v, B, and θ, do not change as the particle moves. Even as the particle's velocity changes direction, the angle θ between the velocity and magnetic field stays constant at $90°$.

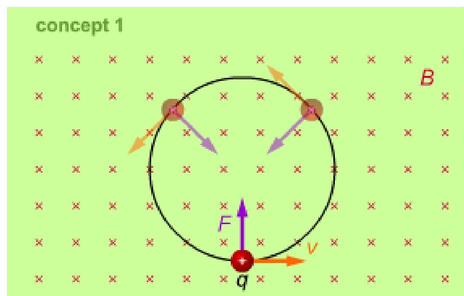
If you recall your studies of uniform circular motion and centripetal forces, this may all sound familiar. A constant force that points toward the center causes uniform circular motion. In other words, the force exerted by the magnetic field is a centripetal force.

An interesting aspect of this kind of motion is that the field does no work on the particle. One way to conclude this is by noting that the particle's energy does not change. Its speed is constant (which means its KE is constant) and its PE does not change in this uniform field. Since there is no change in energy, no work occurs. This is quite distinct from the situation of a charged particle in an electric field: Electric fields can do work on charged particles.

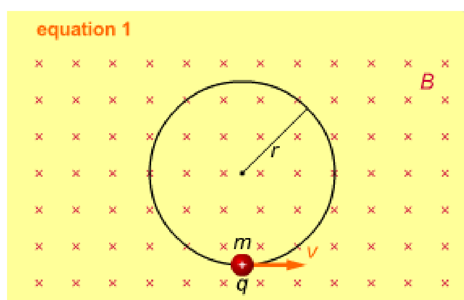
Now we summarize the equations on the right. In Equation 1, we take the centripetal force in uniform circular motion to be the magnetic force, and use Newton's second law to state that this must equal the mass of the particle times its centripetal acceleration v^2/r. We then solve for the radius r.

We can state other equations that further describe the motion of the particle. First, we calculate the period T of the particle's motion: The result is shown in Equation 2. We derive this equation by starting with the equation for the period of an object in circular motion (the circumference divided by the speed). Then we substitute the formula for the speed of the particle in the magnetic field (obtained by solving the radius equation in Equation 1 for v).

Notice the interesting fact that the period is not a function of the speed, but only of the mass and charge of the particle, as well as the strength of the magnetic field. A faster



Circular motion
Velocity perpendicular to magnetic field
· Force perpendicular to velocity
· Magnitude of force constant
· Circular motion results



Radius
$|q|vB = ma = m(v^2/r)$, so

$$r = \frac{mv}{|q|B}$$

r = radius, m = mass

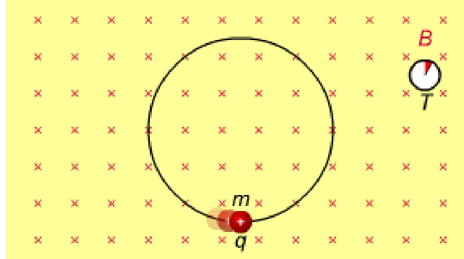
moving particle moves in a circle of greater circumference, but the period does not change.

We have also included in Equation 3 the equations for the frequency f and the angular frequency ω of the particle's motion. These can be derived using the relationship of period and frequency, and of frequency and angular frequency.

Without discussing it further in this section, we do note that if the velocity of the charged particle is not perpendicular to the field, but has a component parallel to the field, then it will move in helical motion. If you would like to observe helical motion, go to the simulation in the introduction to this chapter, move the particle's velocity vector to a direction that is not perpendicular to the field, press GO, and use the viewing angle slider to watch the resulting motion from various vantage points.

v = speed, q = charge
B = magnetic field strength

equation 2



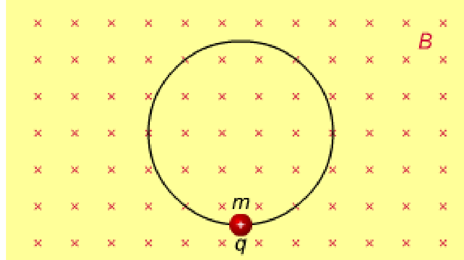
Period

$T = \frac{2\pi r}{v}$, and $v = \frac{|q|rB}{m}$ so

$$\boxed{T = \frac{2\pi m}{|q|B}}$$

T = period

equation 3

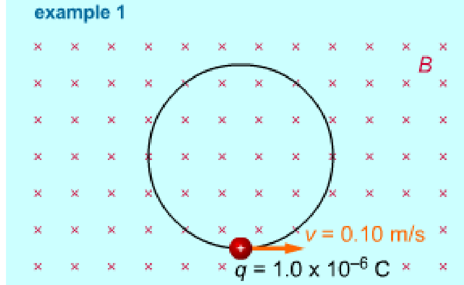


Frequency, angular frequency

$$\boxed{f = \frac{|q|B}{2\pi m}, \quad \omega = \frac{|q|B}{m}}$$

f = frequency, ω = angular frequency

example 1



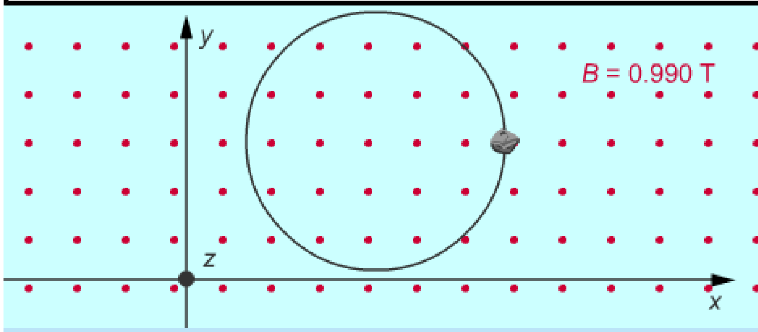
m = 2.5×10⁻⁷ kg and B = 0.50 T. What is the angular frequency of the particle's motion?

$$\omega = \frac{|q|B}{m}$$

$$\omega = \frac{(1.0\times 10^{-6}\text{ C})(0.50\text{ T})}{(2.5\times 10^{-7}\text{ kg})}$$

$\omega = 2.0$ rad/s

30.14 - Interactive checkpoint: a circular path

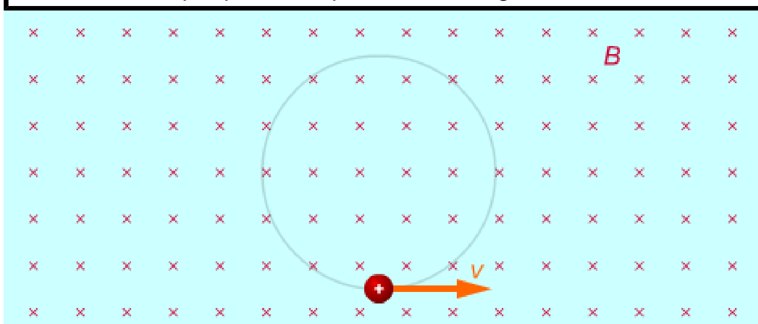


A dust particle has a mass of 5.70×10^{-20} kg and a net charge of 4.80×10^{-19} C. It is observed to travel in a circular path of radius 1.45 cm in the magnetic field shown. At the position shown, state the particle's velocity as a vector in rectangular notation.

Answer:

v = (____ , ____ , ____) m/s

30.15 - Sample problem: proton in a magnetic field



This proton is moving in a circular path in a magnetic field. Its angular velocity is 2.00×10^{7} rad/s.

What is the strength of the magnetic field?

Variables

angular velocity	$\omega = 2.00\times10^{7}$ rad/s
speed	v
radius of circular path	r
magnetic field strength	B
mass of proton	$m_p = 1.67\times10^{-27}$ kg
charge of proton	$q = 1.60\times10^{-19}$ C

What is the strategy?

1. Find the speed of the circling proton in terms of the radius and angular velocity of its motion.
2. State the equation for the radius of the circular motion of a charged particle in a magnetic field. For the particle's speed, use the expression from strategy step 1.
3. Solve for the magnetic field and evaluate.

Physics principles and equations

The relationship between angular velocity and tangential speed is

$$v = r\omega$$

The radius of a particle's circular motion in a magnetic field is

$$r = \frac{mv}{|q|B}$$

Step-by-step solution

Step	Reason
1. $v = r\omega$	angular velocity and speed
2. $r = \dfrac{m_p v}{\lvert q \rvert B}$	radius of circular motion in a magnetic field
3. $r = \dfrac{m_p r\omega}{\lvert q \rvert B}$	substitute equation 1 into equation 2
4. $B = \dfrac{m_p \omega}{\lvert q \rvert}$	simplify and solve for B
5. $B = \dfrac{(1.67\times10^{-27}\text{ kg})(2.00\times10^{7}\text{ rad/s})}{1.60\times10^{-19}\text{ C}}$ $B = 0.209\text{ T}$	evaluate

30.16 - Physics at work: mass spectrometer

Mass spectrometers are used to determine the masses of atoms or molecules, or their relative abundance in a sample. They are used in a range of settings from surgery (to determine the mixture of gases in a patient's lungs) to space missions (to analyze the atmospheres or soils of planets and other celestial bodies). Chemists also frequently use them to analyze materials.

A conceptual diagram of a mass spectrometer is shown to the right. To use the device, the substance being analyzed is first vaporized if it is not already a gas. It is then ionized: An electron is removed from each particle so that it has a net positive charge equal to $+e$. The ionized particles are then accelerated across a potential difference between two charged plates.

The particles all have the same charge, and the experimenter keeps the potential difference between the plates, and the resulting electric field strength, at a constant value. This means the force exerted on various particles by the electric field does not change; their accelerations depend only on their masses.

Each moving, charged particle passes through an entry port into a uniform magnetic field. The field is perpendicular to the velocity of the particle, which means the particle will move in a circular path. The magnetic field is kept constant so that the radius of the path is a function of the particle's speed and mass.

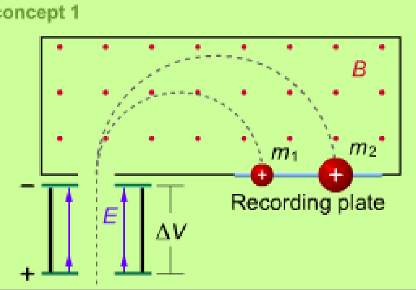
Each particle traverses a circular path and the point where it strikes the sensing plate of the device is recorded. The radius of the path is one-half the distance from the entry port. The particle's mass can be determined using the equation shown to the right.

Determining the mass of a particle (Derivation). In the following derivation we assume that the ion has a charge of $+e$, so that it is accelerated from the positively charged plate to the negatively charged plate, and then curves to the right after it enters the magnetic field, which is directed out of the screen toward you.

Variables

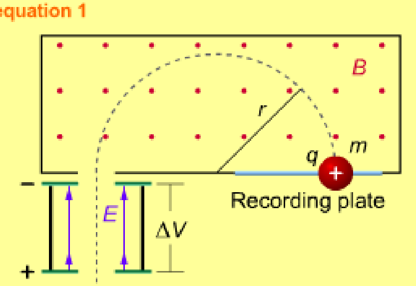
change in kinetic energy	ΔKE
mass of ion	m
accelerated speed of ion	v
radius of circle in magnetic field	r
charge of positive ion	$q = +1.6\times10^{-19}\text{ C}$
magnetic field strength	B
change in potential energy	ΔPE
potential difference across plates	ΔV

concept 1



Mass spectrometer
Device that measures mass of particles
Electric field accelerates ions
Radius measured to determine mass

equation 1



Mass of a particle

$$m = -\frac{qr^2B^2}{2\Delta V}$$

m = mass, q = charge
r = radius
B = magnetic field strength
ΔV = potential difference

Strategy

1. Express the change in the **kinetic** energy of the ion during the acceleration phase in terms of its mass and quantities that can be controlled or measured by the experimenter. By "acceleration phase," we mean the linear acceleration in the electric field.
2. Express the change in the **potential** energy of the ion during the acceleration phase in terms of some of the same quantities, including the potential difference between the plates.
3. Use the conservation of energy to relate kinetic energy to potential energy, and solve for the unknown mass of the charged particle.

Physics principles and equations

We use the definition of kinetic energy.

$$KE = \tfrac{1}{2}mv^2$$

The radius of the circular path of a charged particle moving perpendicular to the magnetic field is

$$r = mv/qB$$

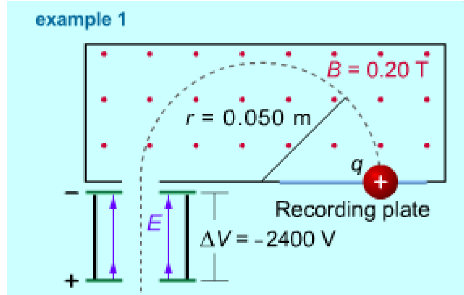
The potential difference equals the change in PE per unit charge.

$$\Delta V = \Delta PE / q$$

The principle of the conservation of energy applied to mechanical energy states that

$$\Delta KE + \Delta PE = 0$$

Since the magnetic field changes the particle's direction but not its speed, its KE is not changed by the magnetic field.



This mass spectrometer is testing a hydrogen molecule (H_2^+) having charge +e. What is its mass?

$q = 1.6\times10^{-19}$ C

$$m = -\frac{qr^2B^2}{2\Delta V}$$

$$m = -\frac{(1.6\times10^{-19}\,\text{C})(0.050\,\text{m})^2(0.20\,\text{T})^2}{2(-2400\ \text{V})}$$

$m = 3.3\times10^{-27}$ kg

Step-by-step derivation

In the first stage of the derivation we write ΔKE in terms of the mass m and the speed v of an ion as it leaves the electric field and enters the magnetic field. Neither m nor v can be directly observed. We use another equation to replace v by quantities that can be controlled or observed in the laboratory.

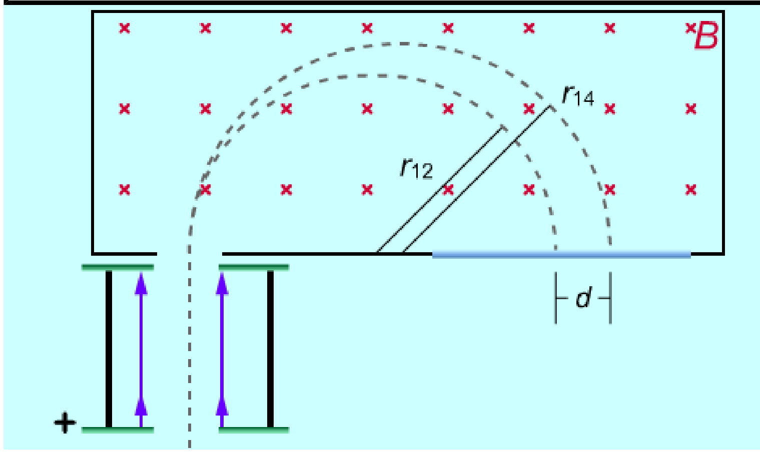
Step	Reason
1. $\Delta KE = \tfrac{1}{2}mv^2$	change in KE for particle starting at rest
2. $r = \dfrac{mv}{qB}$	radius of positive ion's path in magnetic field
3. $v = \dfrac{rqB}{m}$	solve equation 2 for v
4. $\Delta KE = \tfrac{1}{2}m\dfrac{r^2q^2B^2}{m^2}$	substitute equation 3 into equation 1
5. $\Delta KE = \dfrac{r^2q^2B^2}{2m}$	simplify

Now we write ΔPE in terms of q and the potential difference ΔV across the accelerating plates. Finally, we write an equation for ΔKE and ΔPE based on the conservation of energy, substitute the expression found above, and solve for the unknown mass m of the charged particle.

Step	Reason
6. $\Delta PE = q\Delta V$	change in potential energy
7. $\Delta KE = -\Delta PE$	conservation of energy
8. $\dfrac{r^2q^2B^2}{2m} = -q\Delta V$	substitute equations 5 and 6 into equation 7
9. $m = -\dfrac{qr^2B^2}{2\Delta V}$	solve for m

Although we derived the mass equation for a positively charged particle, we could equally well do so for a negatively charged particle (the charged plates in the mass spectrometer would have to be reversed to accelerate it in the correct direction).

30.17 - Sample problem: carbon isotopes



In radiocarbon dating, a mass spectrometer with a uniform magnetic field of 0.0300 T is used to measure the relative abundances of different carbon isotopes. Carbon ions with a charge +*e* enter the device at 2.00×10^4 m/s, travel semi-circular paths, and strike the recording plate.

What distance separates the ^{12}C atoms from the ^{14}C atoms on the plate?

One important application of mass spectroscopy is radiocarbon dating. Archaeologists and paleontologists use this technique to determine the approximate age of organic materials by looking at the ratio of carbon isotopes in a sample.

Variables

speed	$v = 2.00\times10^4$ m/s
magnetic field strength	$B = 0.0300$ T
^{12}C mass	$m_{12} = 1.99\times10^{-26}$ kg
^{14}C mass	$m_{14} = 2.32\times10^{-26}$ kg
ion charge	$q = 1.60\times10^{-19}$ C
^{12}C path radius	r_{12}
^{14}C path radius	r_{14}
isotope separation	d

What is the strategy?

1. State the circular path radius for each ion in terms of the given quantities.
2. Subtract the diameter of one from the other to determine d.

Physics principles and equations

The equation for the path radius of a moving particle in a magnetic field is

$$r = \frac{mv}{|q|B}$$

Step-by-step solution

We begin by stating the path radius for each isotope.

Step	Reason
1. $r = \dfrac{mv}{\lvert q\rvert B}$	path radius of particle in magnetic field
2. $r_{12} = \dfrac{m_{12}v}{\lvert q\rvert B}$ $r_{14} = \dfrac{m_{14}v}{\lvert q\rvert B}$	radii of isotope paths

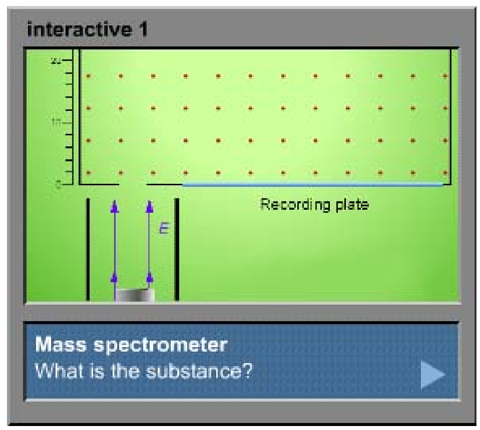
Looking at the illustration tells us the relationship between the radii and the isotope separation: d is the difference between the diameters of the isotope paths, and the diameter of a circle is twice its radius.

Step	Reason
3. $d = 2r_{14} - 2r_{12}$	inspection
4. $d = \frac{2m_{14}v}{\lvert q \rvert B} - \frac{2m_{12}v}{\lvert q \rvert B}$	substitute equation 2 into equation 3
5. $d = \frac{2v}{\lvert q \rvert B}(m_{14} - m_{12})$	factor
6. $d = \frac{2(2.00 \times 10^{4} \frac{\text{m}}{\text{s}})(2.32 \times 10^{-26}\,\text{kg} - 1.99 \times 10^{-26}\,\text{kg})}{(1.60 \times 10^{-19}\,\text{C})(0.0300\,\text{T})}$ $d = 0.0275\,\text{m}$	evaluate

The separation distance is 2.75 cm, or a little over an inch. By measuring the relative abundance of the ^{12}C and ^{14}C isotopes from a sample that arrive at the two locations, scientists are able to deduce the age of the sample.

30.18 - Interactive problem: mass spectrometer analysis

You will use this simulation of a mass spectrometer to determine what substance it is analyzing. You select a velocity and fire a positive ion (charge $+e$) into a magnetic field. It will move along a circular path and strike a recording plate. You can see, and the simulation will tell you with more precision, the radius of its path. You will use that information, the strength of the magnetic field (0.62 T) and the ion's speed to determine its mass.



In the simulation, you will see a drop-down list labeled "Substance is:" It has a list of candidates for the mystery ion, complete with their masses. Determine the mass from your experiment, and then select from among the substances. The simulation will tell you whether you are right or wrong. Be careful with the units: The radius is stated in centimeters.

If you have trouble determining the mass, review the section on the circular motion of particles in magnetic fields, which stated an equation that you can solve to find the mass of a particle based on quantities that you can control or measure in this simulation.

30.19 - Helical particle motion in magnetic fields

In order for a charged particle to move in a circular path in a uniform magnetic field, it must enter the field with a perpendicular velocity. But what if its velocity has a component parallel to the field?

The result is shown to the right: It is called *helical* motion. The particle traces out circles that wind upward (or downward) in a fashion similar to the motion of a car navigating the "corkscrew" ramps found in many multistory parking garages. The particle moves both in circles and up or down at the same time. You can use the interactive simulation in a following section to observe helical motion.

Aurora Borealis display in the northern sky.

To explain why helical motion occurs, we need to decompose the velocity into components perpendicular and parallel to the magnetic field. The two diagrams in Concept 1 do this. In the left-hand diagram, the magnetic field is viewed obliquely (but not fully parallel to the surface of the screen) and you can see both components of the particle's velocity.

The component $v_{\text{perpendicular}}$ is perpendicular to the magnetic field. This component accounts for the circular motion of the particle. The other component of the velocity, v_{parallel}, is parallel to the magnetic field. Since there is no magnetic force exerted on a charge moving parallel to a magnetic field, this velocity component does not change. It accounts for the constant upward or downward motion of the particle.

The result of the forces exerted on the particle is that the particle moves in circles in a plane perpendicular to the magnetic field and at a constant speed in a direction parallel to the magnetic field. The sum of these motions is helical motion.

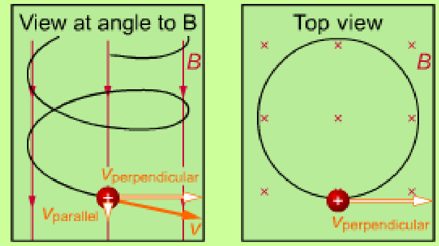
As the particle moves in a helical path, the vertical spacing between the loops of the helix, known as the *pitch*, remains constant since the vertical velocity component does not change.

Helical motion can arise in a nonuniform magnetic field, as well. A magnetic field that is stronger at its outer edges can cause a particle to become trapped in a *magnetic bottle*. The particle moves in a helical fashion, spiraling up and down inside the "bottle." Physicists construct bottles like this as three-dimensional systems that can contain charged particles indefinitely.

The Earth creates a magnetic bottle of this type. Its magnetic field is stronger near the poles. Electrons and protons are trapped in the bottle created by the Earth. They oscillate back and forth over a short distance every few seconds, resulting in what are called the *Van Allen radiation belts*.

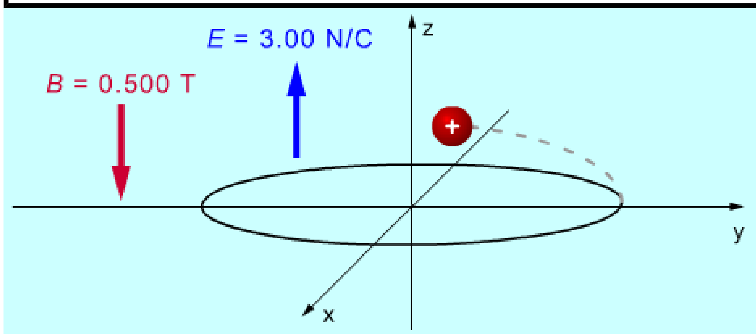
Such belts, as well as solar flares, are responsible for *auroras*, the glorious bands of light visible in the sky at high latitudes at certain times of the year. The auroras result from solar flares that shoot ionized particles, primarily electrons and protons, into the Earth's atmosphere. These particles get trapped in the Van Allen belts. As the particles collide with oxygen and nitrogen molecules from the atmosphere, they emit green and pink light respectively. From a great distance, you may perceive a faint aurora as white light.

concept 1



Helical motion
Velocity components perpendicular, parallel to field
· Perp: force causes circular motion
· Parallel: force zero, $v_{parallel}$ constant
Helical motion results

30.20 - Sample problem: spiraling proton



A proton is orbiting at 335 m/s in the *xy* plane through a 0.500 T magnetic field. When the proton's *y* coordinate is at its maximum, a 3.00 N/C electric field is turned on.

What is the final acceleration vector of the proton after it travels another $\pi/2$ rad around its circular *xy* path?

The magnetic field is directed straight down, and the electric field is directed straight up, as shown in the diagram above. Although the proton has previously been moving in the xy plane only, when the electric field is turned on it will begin to move in the positive z direction as well, exhibiting helical motion.

Variables

magnetic field strength	$B = 0.500$ T
electric field strength	$E = 3.00$ N/C
charge of proton	$q = 1.60\times10^{-19}$ C
mass of proton	$m_p = 1.67\times10^{-27}$ kg
speed of particle in *xy* plane	$v = 335$ m/s
radius of path in *xy* plane	r
electric force on particle	F_E
final acceleration in *xy* plane	$\mathbf{a}_{xy}$
final acceleration along *z* axis	$\mathbf{a}_z$
final acceleration in space	$\mathbf{a}$

What is the strategy?

1. The magnetic field affects the electron's motion in the xy plane only.
2. The electric field affects the motion in the z direction only.
3. Add the acceleration vectors of the proton due to these two fields to get a vector that describes the acceleration in all three dimensions.

Physics principles and equations

A charged particle moves with uniform circular motion in the plane perpendicular to a magnetic field. The radius of its circular path is

$$r = \frac{mv}{|q|B}$$

In uniform circular motion, the centripetal acceleration is directed toward the center of the circle. Its magnitude is

$$a = \frac{v^2}{r}$$

Newton's second law

$$F = ma$$

The relationship between electric field and force is

$$F = qE$$

Step-by-step solution

We start by analyzing the proton's motion in the xy plane, which is due to the magnetic field only.

Step	Reason
1. $a_{xy} = \frac{v^2}{r}$	centripetal acceleration
2. $r = \frac{m_p v}{\lvert q \rvert B}$	path radius of particle in magnetic field
3. $a_{xy} = \frac{v \lvert q \rvert B}{m_p}$	substitute equation 2 into equation 1
4. $a_{xy} = \frac{(335\ \frac{m}{s})(1.60\times 10^{-19}\ C)(0.500\ T)}{1.67\times 10^{-27}\ kg}$ $a_{xy} = 1.60\times 10^{10}\frac{m}{s^2}$ in positive x direction	evaluate

Now we determine the proton's acceleration in the z direction, which is due to the electric field only.

Step	Reason
5. $a_z = F/m_p$	Newton's second law
6. $F_E = qE$	definition of electric field
7. $a_z = qE/m_p$	substitute equation 6 into equation 5
8. $a_z = \frac{(1.60\times 10^{-19}\ C)(3.00\ N/C)}{(1.67\times 10^{-27}\ kg)}$ $a_z = 2.87\times 10^8\ m/s^2$ in positive z direction	evaluate

Finally we add the two component acceleration vectors to find the final acceleration vector in three-dimensional space.

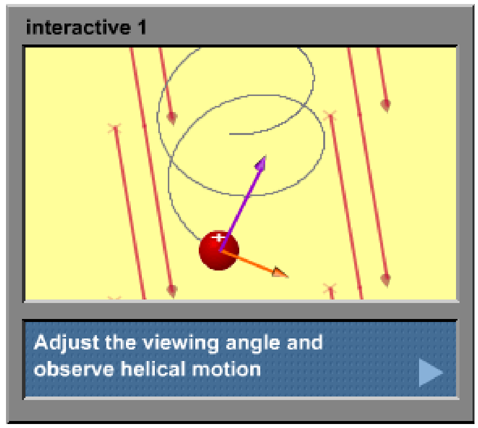
Step	Reason
9. $\mathbf{a} = \mathbf{a}_{xy} + \mathbf{a}_z$	vector equals sum of its components
10. $\mathbf{a} = (1.60\times 10^{10}\frac{m}{s^2},\ 0,\ 2.87\times 10^8\frac{m}{s^2})$	evaluate

30.21 - Interactive problem: helical particle motion

In this simulation you will have a chance to observe a charged particle following a helical path through a magnetic field.

If you use the initial velocity provided by the simulation, which is perpendicular to the field, the particle will move in a circle. You can cause its path to be a helix by supplying a component to its initial velocity that is parallel to the field.

To do this, first make sure the magnetic field lines are viewed as pointing straight down. Now you may drag the tip of the velocity vector arrow to set the initial speed and direction of the particle. Once you have the initial velocity you want, change the viewing angle by moving the slider provided for this purpose to a position near the middle of its range. Press GO to observe the helical motion of the particle.



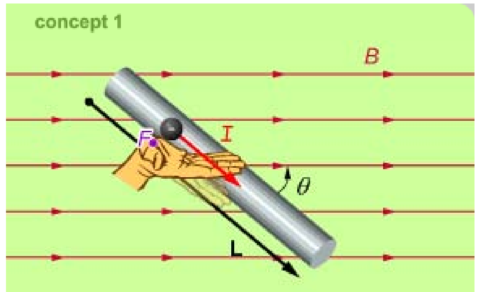
30.22 - Magnetic force on a current-carrying wire

A magnetic field exerts a force on a wire carrying a current. Since the moving charges in the wire − electrons in this case − cannot escape from it, the wire as a whole will react to the magnetic force on them, just as a large net full of helium balloons will rise due to the balloons' individual buoyancies.

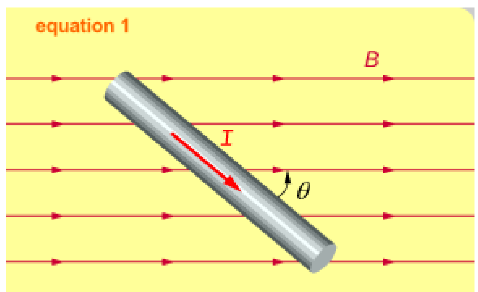
At the right, we depict a configuration that shows how to determine the magnetic force exerted by a uniform magnetic field on a straight, current-carrying wire. We show only a segment of the wire, and not the entire circuit loop that allows current to flow.

The right-hand rule can be used to determine the direction of the force. Since the current shown is conventional (positive), the thumb points in the direction of the force when the fingers wrap from the direction of current flow to the magnetic field. (If the current were shown as flowing electrons, the thumb would point in the direction opposite to the current.) In this case the force points out of the screen, toward you.

To calculate the magnitude of the force exerted on a given length of wire, use the second equation shown in Equation 1. The amount of force increases with the amount of current, the length of the wire and the strength of the magnetic field. Since more current means either more electrons flowing, or the electrons moving faster, this relationship follows from the equation for force on a single charge, $F = qvB \sin \theta$ (or to state it using cross-product notation, $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$). A greater length of wire will also experience more force since it contains more moving charge.



Magnetic force on a wire
Uniform magnetic field exerts force on current-carrying wire
Proportional to current, length of wire
Direction of force found with right-hand rule



Magnetic force on a wire

$\mathbf{F} = I\mathbf{L} \times \mathbf{B}$

$F = ILB \sin \theta$

$\mathbf{F}$ = force, I = current
$\mathbf{L}$ = directed length of wire segment
$\mathbf{B}$ = magnetic field
θ = angle between wire and field

Variables

We list below the variables that are used in the derivation, but which do not already appear in Equation 1. The vector $\mathbf{L}$ appearing in the equation is the "directed length" of the wire segment. That is, $\mathbf{L}$ is parallel to the segment, in the direction of the current, and its magnitude L equals the length of the segment.

amount of free charge in wire segment	Q
velocity of a charge carrier in segment	$\mathbf{v}$
time for carrier to move through segment	Δt

Strategy

1. Use the definition of the current flowing through the segment to find that $I\mathbf{L} = q\mathbf{v}$.
2. Substitute this equation into the cross-product formula for the force exerted on a moving charge by a magnetic field to get Equation 1.

Physics principles and equations

The force exerted on a moving charged particle by a magnetic field is,

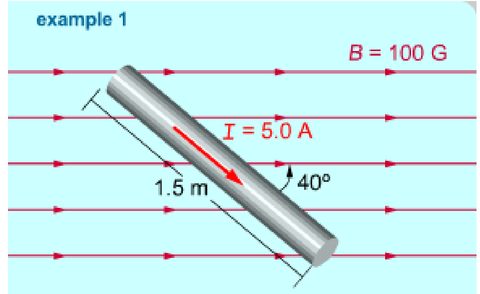
$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}$$

Step-by-step derivation

We use the definition of the current flowing through the wire to obtain an equation relating $I\mathbf{L}$ and $q\mathbf{v}$. This enables us to make a substitution into the equation for the force exerted by a magnetic field on a moving charge.

Step	Reason
1. $I = q/\Delta t$	definition of current
2. $I\mathbf{L} = q\mathbf{L}/\Delta t$	multiply both sides by $\mathbf{L}$
3. $I\mathbf{L} = q\mathbf{v}$	$\mathbf{L}/\Delta t$ equals velocity of charge carrier
4. $\mathbf{F}_B = q\mathbf{v} \times \mathbf{B}$	equation for force on charged particle
5. $\mathbf{F}_B = I\mathbf{L} \times \mathbf{B}$	substitute equation 3 into equation 4

We state the equation proved above as a cross product. It is the same as saying the amount of force is $F_B = ILB \sin\theta$.



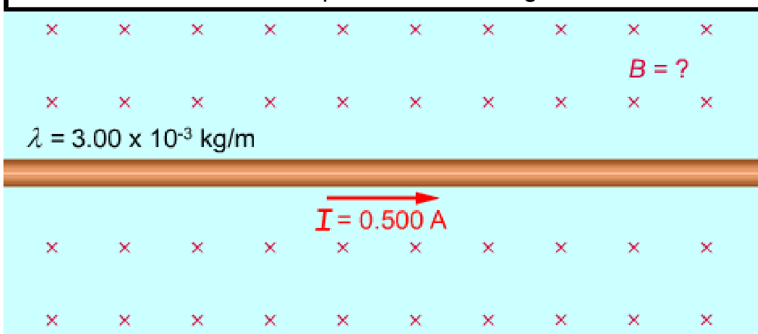
What is the force on the wire?

$F = ILB \sin\theta$

$F = (5.0\text{ A})(1.5\text{ m})(0.010\text{ T})(\sin 40°)$

$\mathbf{F} = 4.8\times10^{-2}$ N straight towards you

30.23 - Interactive checkpoint: the levitating wire



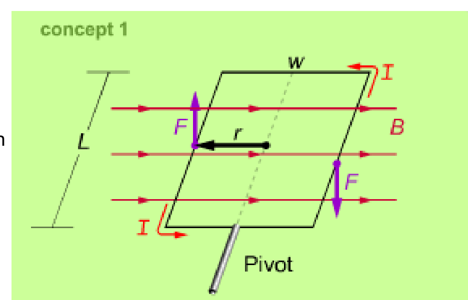
A long horizontal wire with linear mass density $\lambda = 3.00\times10^{-3}$ kg/m carries a current of $I = 0.500$ A. A horizontal magnetic field is applied perpendicularly to the wire. What strength of magnetic field will keep the wire "levitating"? In other words, the magnetic field force equals the force of gravity along any segment of the wire.

Answer:

$B =$ ______ T

30.24 - Torque on a loop of current

Magnetic fields and loops of current-carrying wire are essential components of electric motors. Applications like this make calculating the torque on a loop of wire a useful exercise for engineers and physics students alike.



Torque on current loop

Current loop in magnetic field experiences torque

In Concept 1, you see an illustration of a rectangular loop of wire in a magnetic field. The magnetic field points to the right. The conventional current runs counterclockwise in the loop. We start this discussion by considering the loop when the surface area enclosed by the loop lies along the field, as it does in Concept 1. Next, we will consider what occurs when the loop rotates.

As the diagram reflects, the force is upward on the left side of the loop and downward on the right side. You can verify the direction of the force using the right-hand rule. The forces from the left and right sides of the loop create torque that will cause the loop to rotate clockwise around its pivot axis.

Only the forces on the left and right sides of the loop create a net torque. Combined, the front and back sections experience no net torque. Why? In the initial position shown in Concept 1, there are no forces on the front or back parts of the loop because the current through them runs parallel to the field. As the loop rotates from this position, the field does exert a force on these segments, but in opposing directions along the pivot axis, so there is no net force or torque on these sections of the loop.

Now let's consider the torque in more detail. Torque equals the cross product $\mathbf{r} \times \mathbf{F}$, where $\mathbf{r}$ is a position vector from the pivot axis out to the point on the wire where the force $\mathbf{F}$ is applied. In the initial position, since the force is perpendicular to $\mathbf{r}$, the amount of torque equals rF.

The magnitude of the position vector r equals one half the width of the loop, $\frac{1}{2}w$, so the torque on one side of the loop equals $\frac{1}{2}wF$. (Sometimes r is referred to as "the radius" of the loop.) Because there are two equal forces applying torque on opposite sides of the loop, the magnitude of the total torque equals wF.

The force F on a wire of length L carrying a current I perpendicular to a magnetic field $\mathbf{B}$ equals ILB. Multiplying this expression for the force by w, as above, and rearranging yields $IwLB$. We can simplify this expression: Since the width times the length of the loop (wL) equals its area A, the total torque on the loop around its pivot axis is IAB.

In the illustration for Concept 2, we show the wire loop from a different point of view, this time from behind the pivot rod. The loop has been rotated approximately $30°$ from its initial position. The current on the left runs toward you (as indicated by the red arrow point) and the current on the right is directed away.

The current in the left and right sides still flows perpendicularly to the magnetic field. However, since the position vector **r** and the force vector **F** are no longer perpendicular, we must factor in this angle, using the sine function, to calculate the torque.

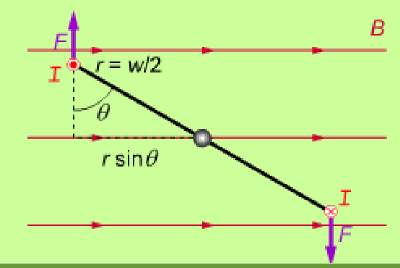
We will factor in the angle using an area vector. In Equation 1 we have drawn the area vector **A** for the loop. An area vector's length is proportional to the amount A of the surface area, and it is perpendicular to the surface area. Since the angle between **r** and **F** is the same as the angle between **A** and **B**, we can express the torque as $I\mathbf{A} \times \mathbf{B}$.

A quick check confirms this equation: The torque is at its maximum when **A** is perpendicular to the field (the sine of $90°$ is one), that is, when the loop and magnetic field are aligned as shown in Concept 1. The torque is zero when **A** is parallel to the field (the sine of $0°$ is zero). The torque is zero at this orientation because when the plane of the loop is vertical, the two force vectors on the lengthwise sections of the loop each have a moment arm of zero, and there is no torque when **r** equals zero.

The primary component of an electric motor is a coil of wire in a magnetic field. The comments below may give you a sense of why it takes several additional components to turn this configuration into a useful motor. Consider what will happen if angular momentum causes the loop to rotate beyond the straight up-and-down position where it experiences zero torque.

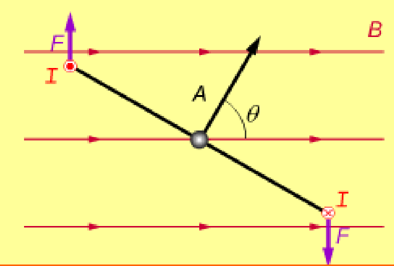
Now the current on the left flows away from you, and the current on the right flows toward you. The force on the left points down, and the force on the right points up. This means the direction of the torque is reversed, and tends to make the loop rotate counterclockwise, back toward its up-and-down position perpendicular to the field. The torque is minor when the coil is near this orientation, so it is not very effective for a motor; the coil will move to this vertical orientation, oscillate back and forth a bit, and stop. However, by reversing the direction of current flow as the coil rotates, a clever engineer could create a torque that would cause the loop to continue to rotate, and indeed, this is how motors function.

concept 2



View behind pivot
Orientation of loop affects torque

equation 1



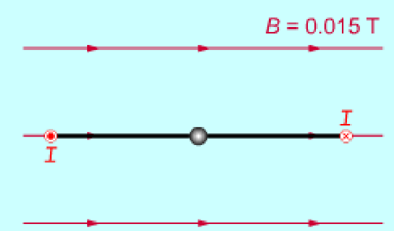
View behind pivot

$$\boldsymbol{\tau} = I\,\mathbf{A} \times \mathbf{B}$$

$$\tau = IAB \sin\theta$$

$\boldsymbol{\tau}$ = torque
I = current, **A** = area vector
B = magnetic field
θ = angle between **A** and **B**

example 1



Twenty amperes flows through a loop 0.50 m by 0.30 m. What is the maximum torque on the loop?

$\tau = IAB \sin\theta$
$A = (0.50 \text{ m})(0.30 \text{ m}) = 0.15 \text{ m}^2$
$\tau = (20 \text{ A})(0.15 \text{ m}^2)(0.015 \text{ T})(1)$
$\tau = 0.045 \text{ N·m}$

30.25 - Magnetic dipole moment

A coil of current-carrying wire, a bar magnet, a compass needle: All will experience a torque when placed in a magnetic field. Like the bar magnet or the needle, the coil is a magnetic dipole. The *magnetic dipole moment* $\boldsymbol{\mu}$ is a vector used to quantify how much torque a magnetic dipole will experience in an external magnetic field.

We will use the example of the loop of wire shown to the right to explain how the magnetic moment is calculated for coils.

The torque on a single loop of wire equals $I\mathbf{A} \times \mathbf{B}$. If there are N equivalent loops of wire, then the torque equals $NI\mathbf{A} \times \mathbf{B}$. The first factor in this cross product defines the magnetic dipole moment, the magnitude of which equals NIA, and whose direction is the same as that of $\mathbf{A}$. The torque on a magnetic dipole then equals $\boldsymbol{\mu} \times \mathbf{B}$.

We state these relationships in Equations 1 and 2. In Equation 1, we define the dipole. To determine the direction of the dipole moment, use a right-hand rule: The fingers wrap in the direction of conventional current flow through the coil; the thumb will point in the direction of the moment. It is parallel to the area vectors of the loops of the coil, that is, perpendicular to the plane that contains the coils. In Equation 2, we show how the torque is calculated.

The magnetic dipole moment can be used to calculate not just the torque on a dipole, but also the magnetic potential energy of the dipole in a magnetic field. The potential energy equals the negative of the dot product of the moment and the field. This relationship is stated in Equation 3.

By convention, the potential energy of a dipole is zero when the dipole moment is perpendicular to the field. The magnetic potential energy is at its highest value when the dipole moment is aligned exactly opposite (at a $180°$ angle) to the magnetic field, and at its lowest (a negative value) when the moment is aligned with the magnetic field. When the moment is "balanced" oppositely to the magnetic field, it will rotate toward alignment if the coil is given the slightest push in either direction, an indication that it has PE. The dipole orients with the magnetic field, so it can be used like the magnetic needle in a compass.

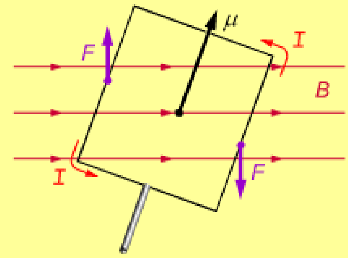
The change in PE as the coil rotates from one position to another equals the work done on the coil by an external agent to cause it to rotate between those positions, from one stationary position to another. Since work also depends on the product of torque and angular displacement, the computation of work, using calculus, can be used to derive the stated equation for PE from the formula for torque.

The magnetic moment of a bar magnet points from the magnet's south pole to its north pole. This is shown in Equation 4.

Since a current-carrying coil of wire has a magnetic dipole moment just like that of a bar magnet, it can be thought of as a magnet in its own right, with its magnetic moment pointing in the direction of its magnetic north pole. In fact, current-carrying coils are often called *electromagnets*. They may be wrapped around a core consisting of a material like iron in order to create an enhanced magnetic field in the iron.

Some atoms also have magnetic dipole moments. An external magnetic field will exert torque on them, causing them to align with it. This is the basis for the *MRI* (*magnetic resonance imaging*) machine, a technology used in hospitals to image the interior of the human body.

equation 1



Magnetic moment of a coil

$\boldsymbol{\mu}$ is magnetic dipole moment

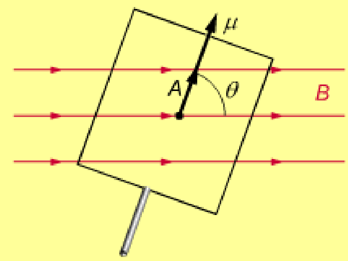
Magnitude: $\mu = NIA$

N = number of loops, I = current
A = area of one loop

Direction: right hand rule

Units: ampere·meter^2, joule/tesla

equation 2



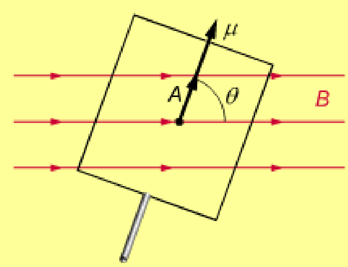
Torque

$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$

$\tau = \mu B \sin\theta$

τ = torque, μ = dipole moment
B = magnetic field strength
θ = angle between moment and field

equation 3



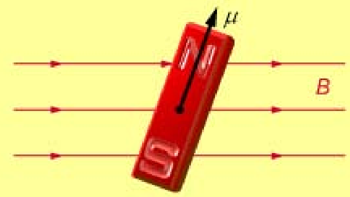
Potential energy

$PE = -\,\boldsymbol{\mu} \cdot \mathbf{B}$

$PE = -\,\mu B \cos\theta$

PE = potential energy

equation 4

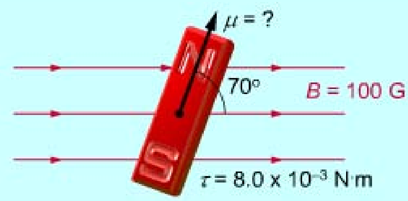


Bar magnet

μ is magnetic dipole moment

Direction: from south to north pole

example 1



What is the dipole moment of the magnet?

$\tau = \mu B \sin\theta$

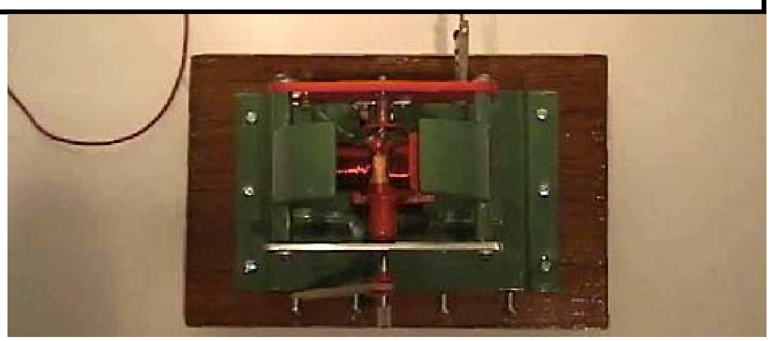
$$\mu = \frac{\tau}{B\sin\theta}$$

$$\mu = \frac{(8.0\times 10^{-3}\ \mathrm{N\cdot m})}{(0.010\ \mathrm{T})(\sin 70°)}$$

$\mu = 0.85$ J/T

30.26 - Physics at work: direct current electric motor

Direct current motors power a range of familiar everyday devices, from hairdryer fans to handheld electric drills. In these appliances, a few fundamental physics principles and clever engineering combine to create a motor that yields a constant amount of torque.



DC electric motor. The sparking copper commutator is visible between the rotors and the red plate at the top of the image.

At its simplest, a direct current motor consists of a current-carrying wire coil, wrapped around a metal armature, inside a uniform magnetic field. The external field creates a torque on the coil when current flows through it. That torque is used to rotate something: a fan, a drill bit, the blades of a blender and so forth.

However, a motor has two requirements that this configuration alone does not meet. First, it needs to rotate continuously in the same direction. The simple wire coil with a current will not do this: It will rotate in one direction until the field exerts no net torque on it, continue on beyond that due to its momentum, and then rotate back.

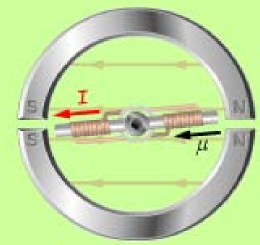
The second requirement for the motor is that it should provide a nearly constant torque. The torque on a simple coil varies as the angle of the coil in the field changes.

In Concept 1, you see a schematic diagram of a simplified direct current motor. Two permanent magnets form the circular outer edge of the motor. A magnetic field is directed from the north poles of these two magnets to the south poles. Inside these magnets is a wire coil, which is connected to a direct current, such as the current from a battery. This assembly is called a *rotor*.

The problem of rotating the rotor in a constant direction is solved with a commutator. A *commutator* consists of two sets of contacts that supply current to the coil. During one-half of a rotation, the coil is in contact with one set and the current flows in one direction. During the other half turn, the coil is in contact with the other set and the current flows in the opposite direction. In Concept 2, you can see that the current keeps reversing direction, which alters the direction of the coil's magnetic dipole moment. This reversal of current causes the rotor to keep experiencing a counterclockwise torque.

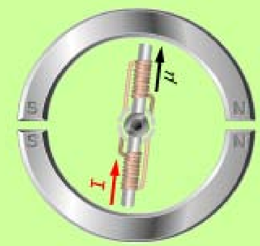
The problem of supplying constant torque is addressed with multiple rotors as you see in Concept 3. At any moment in time, current flows through two of the rotors, which are at different angles to the permanent field. Although the torque on any one rotor varies as it rotates, the sum of the torques on all three rotors stays nearly constant. This combination of current-reversing commutators with multiple rotors enables the motor to provide a steady source of torque.

concept 1



Electric motor
Rotor: bar surrounded by wire coil
Current flows through coil
Current creates magnetic moment
External field exerts torque on coil

concept 2

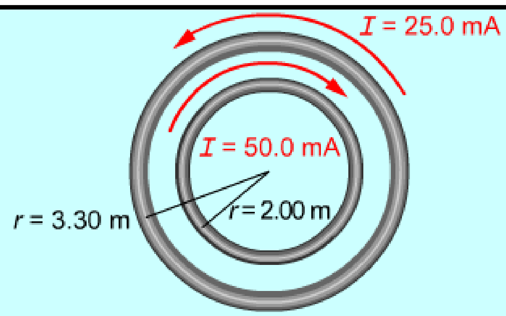


Commutator
Changes direction of current
Reverses magnetic moment

concept 3

Rotors with multiple arms
More arms = smoother torque

30.27 - Sample problem: magnetic dipole moment



A wire loop of radius 2.00 m carries a clockwise current of 50.0 mA.
A concentric loop of radius 3.30 m carries a counterclockwise current of 25.0 mA.

What is the magnetic dipole moment **μ** of the system?

This problem requires that you find the magnetic dipole moments associated with **circular** loops of current. We have already developed formulas for the magnetic moment and torque on a **rectangular** loop of current, based on the area A of the loop. In fact the moment and torque depend only on the area of the loop, and not on its shape.

(To see this, break up any current-carrying loop into a mosaic of tiny separate rectangular loops, each carrying the same current in the same wise. All of the "interior" currents of adjacent rectangles will cancel, their total area equals the area of the original loop, and the two arrangements are equivalent.) This means that you can use the formula for any loops, including the circular ones here.

Variables

	Inner loop	Outer loop
radius of loop	$r_i = 2.00$ m	$r_o = 3.30$ m
area of loop	A_i	A_o
current through loop	$I_i = 0.0500$ A	$I_o = 0.0250$ A
dipole moment of loop	$\boldsymbol{\mu}_i$	$\boldsymbol{\mu}_o$
total magnetic dipole moment	$\boldsymbol{\mu}$	

What is the strategy?

1. Express the magnetic dipole moment of each loop in terms of known quantities.
2. Add the two vectors together to get the total magnetic dipole moment, invoking the principle of superposition.

Physics principles and equations

The equation for the magnitude of the magnetic dipole moment of a coil is

$$\mu = NIA$$

The direction of the moment is determined by a right-hand rule.

The principle of superposition means that the magnetic moment of the system of two loops equals the sum of their individual moments.

$$\boldsymbol{\mu} = \boldsymbol{\mu}_i + \boldsymbol{\mu}_o$$

Mathematics principles

We will use the formula for the area of a circle,

$$A = \pi r^2$$

Step-by-step solution

First we develop an equation for the magnitude of the magnetic dipole moment of any circular current-carrying loop of radius r.

Step	Reason
1. $\mu = NIA$ $\mu = IA$	magnetic moment of a coil with one loop
2. $A = \pi r^2$	area of circular loop
3. $\mu = I\pi r^2$	substitute equation 2 into equation 1

The total magnetic dipole moment is the vector sum of the moments of each loop, by the principle of superposition. The direction of each vector is found using a right-hand rule.

Step	Reason
4. $\mu_i = I_i \pi r_i^2$ $\mu_i = (0.0500\text{ A})\pi(2.00\text{ m})^2$	substitute values for inner loop
5. $\boldsymbol{\mu}_i = 0.628$ J/T away from you	evaluate
6. $\mu_o = I_o \pi r_o^2$ $\mu_o = (0.0250\text{ A})\pi(3.30\text{ m})^2$	substitute values for outer loop
7. $\boldsymbol{\mu}_o = 0.855$ J/T toward you	evaluate
8. $\boldsymbol{\mu} = \boldsymbol{\mu}_i + \boldsymbol{\mu}_o$ $\boldsymbol{\mu} = 0.227$ J/T toward you	use superposition to evaluate total moment

30.28 - Gotchas

Compass needles point north because Earth's magnetic north pole is located near the geographic North Pole. No, the north end of the needle on a compass is attracted to the magnetic south pole of the Earth. Compass needles point approximately to the geographic North Pole because the magnetic south pole of the Earth is near its geographic North Pole.

Magnetic force vectors on charged particles point in the same direction as the magnetic field. No, the magnetic force vectors which act on moving, electrically charged particles are perpendicular to the magnetic field, and to the particles' velocities, as well.

Magnetic fields exert a force on all moving electrically charged particles. Almost true. The charges have to be moving for there to be a force, but if they are moving parallel to or opposite to the field, there will be no force. There must be at least some component of the velocity perpendicular to the field for a force to exist.

When I use the right-hand rule for any charged particle moving in a magnetic field, my thumb points in the direction of the magnetic force. No, the right-hand rule gives the direction of the force on a positively charged particle. For negative particles, the thumb points opposite to the direction of the magnetic force.

30.29 - Summary

A magnet is an object that creates magnetic fields and can exert a magnetic force on other magnets or on moving charged particles. Magnets always have two poles, called north and south poles. As with electrical charges, opposite poles attract each other and like poles repel.

The exterior field lines of the magnetic field generated by a magnet are directed from its north pole to its south pole. The symbol for a magnetic field is $\mathbf{B}$, a vector. Magnetic fields are measured in teslas (T). 1 T = 1 N·s / C·m. Magnetic fields − especially weaker ones − are also measured in smaller units called gauss (G). 1 G = 10^{-4} T.

The Earth has its own magnetic field, which is why compasses work on its surface. The Earth's magnetic south pole is near the geographic North Pole. Because the two do not coincide, when using a compass you need to know the declination, the angle between the magnetic and geographic poles at your location. Compasses can also help you estimate your latitude if they are allowed to orient in three dimensions.

A magnetic field exerts a force on a moving charge. The force is perpendicular to both the velocity of the charge and the magnetic field. If you wrap the fingers of your right hand from the velocity vector to the magnetic field vector, your thumb points in the direction of the force on a positive charge. Your thumb is pointing in the direction opposite to the force on a negative charge.

A charged particle that is moving perpendicularly to a uniform magnetic field will move in a circular path. If it has a velocity component parallel to the field, that component will cause helical motion: The particle will move in a circular path (in two dimensions) while moving at a constant velocity in the third dimension.

A device called a mass spectrometer takes advantage of the circular motion caused by a magnetic field to separate moving particles by their mass-to-charge ratios.

The Hall effect is the name given to the fact that charge carriers in a current-carrying wire immersed in a perpendicular magnetic field are forced to one side of the wire, in a third direction perpendicular to both the direction of current flow and the field. The Hall effect shows that the current carriers in some common conductors such as silver or copper are negative. In fact, they are electrons.

Since an electric current consists of moving charges, a current-carrying wire can have a force exerted on it by a magnetic field. The strength of the force is proportional to the magnitude of the current, the length of the wire, the magnetic field strength, and the sine of the angle between the wire and the field.

A loop of current will experience a torque in a magnetic field.

The magnetic dipole moment $\boldsymbol{\mu}$ is a vector that quantifies how much torque a magnetic dipole (like a current loop) will experience in an external magnetic field.

An electric motor takes advantage of the torque exerted by a magnetic field on a loop of current to spin a rotor, which turns a shaft to perform useful work.

Equations

Force on charge moving in B field

$$F = |q|vB\sin\theta$$

Motion of a charge in a B field

$$r = \frac{mv}{|q|B}$$

In a mass spectrometer

$$m = -\frac{qr^2B^2}{2\Delta V}$$

Magnetic force on a wire

$$F = ILB\sin\theta$$

Torque on a loop of current

$$\boldsymbol{\tau} = I\mathbf{A} \times \mathbf{B}$$

$$\tau = IAB\sin\theta$$

Magnetic moment of a coil

$$\mu = NIA$$

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$$

$$\tau = \mu B\sin\theta$$

$$PE = -\boldsymbol{\mu} \cdot \mathbf{B}$$

Chapter 30 Problems

Chapter Assumptions

Elementary charge, $e = 1.60 \times 10^{-19}$ C

Mass of electron, $m_e = 9.11 \times 10^{-31}$ kg

Mass of proton, $m_p = 1.67 \times 10^{-27}$ kg

Unless stated otherwise, use 5.00×10^{-5} T for the strength of the Earth's magnetic field at its surface.

Conceptual Problems

C.1 While you are on a visit to Santa Claus at his workshop in the Far North, he gives you a special compass whose needle can rotate freely in three dimensions. You try it out by standing directly above the nearby magnetic pole of the Earth. Which direction does the north end of the compass needle point? Explain.

i. North
ii. South
iii. East
iv. West
v. Up
vi. Down

C.2 A charged particle moves in a straight line through a region of space. (a) Can there be a nonzero magnetic field in this region of space? (b) If not, explain why not, and if so, explain in what direction(s) it must point.

(a) ○ Yes ○ No
(b)

C.3 Suppose you are "chirally challenged" (like one of the authors of this textbook, and, as it happens, the physicist Heinrich Hertz): All too often, you cannot accurately recall which is right and which is left. Not only do your driving friends refuse to let you "navigate," but you occasionally use your left hand by mistake when attempting to find the direction of the force exerted on a moving proton by a magnetic field. (a) Do you get the right direction anyway? (b) For which of the following particles does your left thumb point in the direction of the force?

(a) ○ Yes ○ No
(b) i. Positron
ii. Neutron
iii. Electron
iv. Neutrino
v. Proton

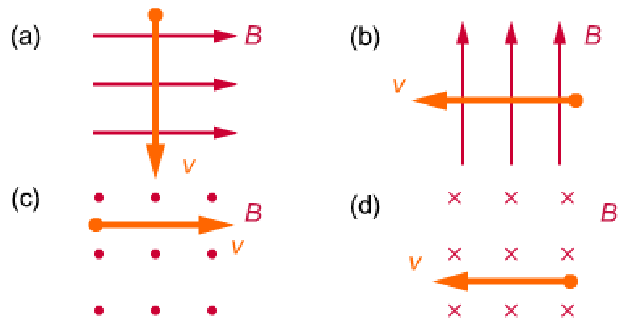
C.4 Can the force from a steady magnetic field cause the speed of a moving charged particle to change? Explain.

○ Yes ○ No

C.5 Can the force from a steady magnetic field cause a stationary charged particle to start moving? Explain.

○ Yes ○ No

C.6 Each of the diagrams (a), (b), (c), and (d) represents a positively charged particle moving at velocity **v** through a magnetic field **B** whose field lines are shown. In each case, tell the direction of the force experienced by the moving particle in the field.



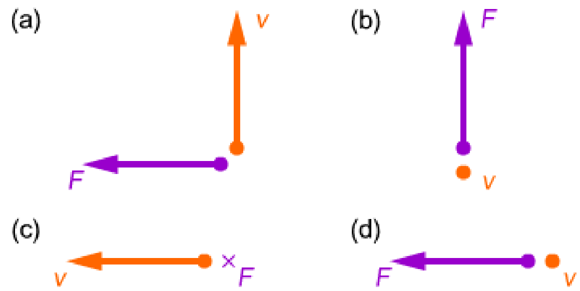
(a) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

(b) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

(c) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

(d) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

C.7 Each of the diagrams (a), (b), (c), and (d) represents a **negatively** charged particle moving at velocity **v** through a magnetic field **B** whose field lines are not shown. In each case the field exerts a force **F** on the particle in the direction shown. In what direction must the field be pointing?



(a) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

(b) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

(c) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

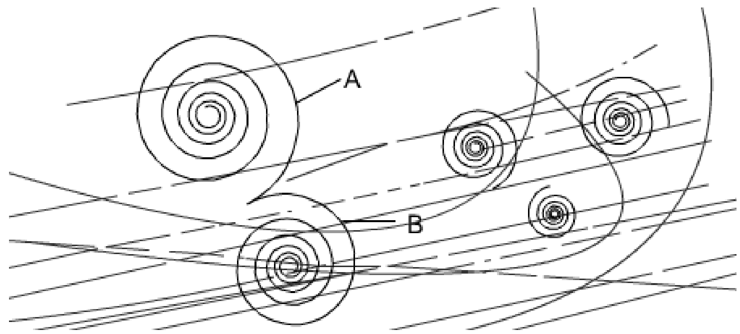
(d) i. Right
ii. Left
iii. Up
iv. Down
v. Toward you
vi. Away from you

C.8 In each part of this problem, you know a charged particle is moving at a constant velocity, and you are asked to draw conclusions about the electric and magnetic fields around the particle. Consider each of the cases independently, and ignore any gravitational effects. In each case, check all of the configurations consistent with the situation, and explain your answer. (a) There is no magnetic field present. (b) There is no electric field present. (c) There is a magnetic field present, with a component perpendicular to the particle's motion.

(a) ☐ There is no electric field
☐ There is an electric field parallel to the velocity
☐ There is an electric field perpendicular to the velocity
☐ None of these are possible

(b) ☐ There is no magnetic field
☐ There is a magnetic field parallel to the velocity
☐ There is a magnetic field perpendicular to the velocity
☐ None of these are possible

(c) ☐ There is no electric field
☐ There is an electric field parallel to the velocity
☐ There is an electric field perpendicular to the velocity
☐ None of these are possible



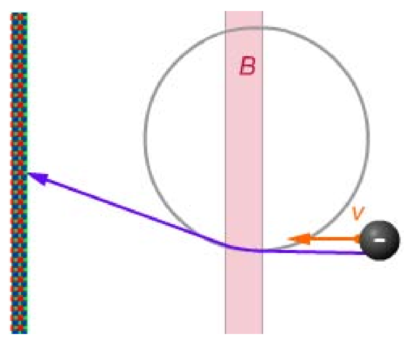
C.9 The picture shows the paths of some neutral and charged particles in a bubble chamber, which contains an undisturbed liquid heated above its boiling point. The idea is that bubbles begin forming on nucleation centers such as surface irregularities, or particles moving through the liquid. These bubbles allow you to see the trails of the particles. Particles enter the chamber from the left. A uniform magnetic field is directed out of the screen toward you. A certain incoming particle (whose trail is not visible) splits into 3 other particles, including those that follow paths A and B. (a) What is the sign of the charge of the particle that follows path A? (b) What is the sign of the charge of the particle that follows path B? (c) Explain your answers.

(a) i. Positive
ii. Negative
iii. Neutral

(b) i. Positive
ii. Negative
iii. Neutral

(c)

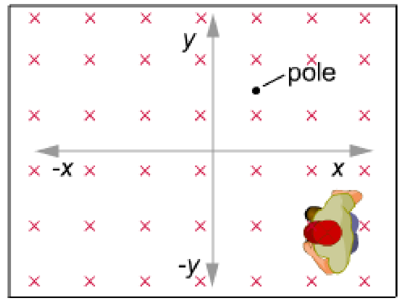
C.10 Some computer monitors are made with cathode ray tubes (CRTs) like a traditional TV screen. The glowing image on a CRT monitor is created by a steerable stream of electrons that travel from the back of the tube to strike tiny phosphors on the inside of the screen surface. When a bar magnet is placed near a CRT computer monitor, it causes severe (and possibly lasting) distortion of the image displayed on the screen. Describe what causes the distortion.



C.11 A focused magnetic field of variable strength is used to "steer" electrons in a television tube (CRT). The electrons emerge in a stream from a "gun" that has accelerated them across a potential difference. A magnetic field can deflect the beam of electrons in various directions. They then travel to the tube's front inside surface on the TV screen, where they strike red, green, or blue phosphors and cause them to glow, creating part of one pixel in a picture display (actual color tubes use three guns, one for each color). Use the orientation of the components as shown in the illustration to answer these questions. (a) In what direction should the magnetic field point to deflect the electron beam upward, toward the top of the tube? (b) A TV tube plays the electron beam over its picture screen in a *raster* pattern, moving it from one side of the screen to the other at the very top of the tube, then all the way across the screen at a lower point, and again and again at still lower points, until the whole screen is painted with sideways sweeps. A second magnetic field of varying strength, which is not shown in the illustration, deflects the beam from side to side at each one of its vertical positions. Suppose a beam is heading straight toward the left in the illustration (it is not deflected vertically). In what direction would the second field have to point to deflect the beam toward you?

i. Right	i. Right
ii. Left	ii. Left
iii. Up	iii. Up
iv. Down	iv. Down
v. Toward you	v. Toward you
vi. Away from you	vi. Away from you

C.12 You are standing in the middle of the recreation room of a nuclear submarine that is lurking in the Arctic Ocean at the magnetic south pole of the earth, where the magnetic field lines point straight down toward the ocean floor, in the negative *z* direction. You can fire positively charged particles from the origin at various speeds in any one of the four directions along the positive or negative *x* and *y* axes. There is a vertical pole in the position shown. In which direction(s) can you fire positively charged particles and have them reach the pole? Assume that the room is large enough that you don't have to worry about hitting the walls.



- ☐ Positive *x*
- ☐ Negative *x*
- ☐ Positive *y*
- ☐ Negative *y*

C.13 A charged particle is moving freely through a uniform field. How can you tell if the field is an electric field or a magnetic field?

C.14 Cosmic "rays", which are charged particles approaching the Earth from space, strike the planet more frequently at the poles than at the equator. Why is this?

C.15 A current-carrying wire is placed in a uniform magnetic field, but there is no magnetic force on it. How is this possible?

C.16 Can one place a current loop in a magnetic field so that it experiences no torque from the field? Explain.

○ Yes ○ No

C.17 The magnetic dipole $\boldsymbol{\mu}$ has magnitude 0.125 J/T and it makes an angle θ with a uniform magnetic field of strength 0.0340 T. When $\theta = 0°$, the dipole is aligned with the field, when $\theta = 90°$, the dipole is perpendicular to the field, and when $\theta = 180°$, the dipole points directly opposite to the field. (a) For which value of θ is the magnetic potential energy of the dipole the greatest (most positive)? (b) For which value of θ is the magnetic potential energy of the dipole the least (most negative)? (c) For which value of θ is the torque on the dipole the greatest? (d) For which nonzero value of θ is the torque on the dipole the least?

(a) __________ °
(b) __________ °
(c) __________ °
(d) __________ °

C.18 Does a uniform magnetic field exert a net force on a bar magnet if the magnet is perpendicular to the field?

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the first interactive problem, and the initial conditions, in this section to answer the following questions. (a) Does the moving particle travel in a straight line or a curve? (b) Is the particle's speed changing? (c) Is the particle accelerating?

(a) ○ Straight line ○ Curve
(b) ○ Yes ○ No
(c) ○ Yes ○ No

0.2 Use the simulation in the first interactive problem, and the initial conditions, in this section to answer the following questions. (a) Are the force and velocity vectors parallel or perpendicular to each other? (b) What is the angle between the force vector and the magnetic field?

(a) ○ Parallel ○ Perpendicular
(b) i. 0 degrees
ii. 45
iii. 90

Section 7 - Magnetic fields and charged particles

7.1 An alpha particle ($q = +3.20\times10^{-19}$ C) moves at a velocity of 2.45×10^4 m/s in the positive x direction through a magnetic field of 0.0775 T that points in the negative z direction. What are (a) the magnitude and (b) the direction of the magnetic force on the particle?

(a) ________ N

(b) i. Positive x direction
ii. Positive y
iii. Positive z
iv. Negative x
v. Negative y
vi. Negative z

7.2 A charged particle moving at a speed of 175 m/s due east, through a magnetic field of 5.00 mT pointing straight up toward the sky, experiences a force of 4.25 μN directed due north. What is the charge on the particle? Be sure to indicate the sign.

________ C

7.3 An electron ($q = -1.60\times10^{-19}$ C) is traveling through the Earth's magnetic field at a location on the Earth's surface where the field has a strength of exactly 5.00×10^{-5} T, directed due north. At a particular instant, a magnetic force of 6.44×10^{-21} N, directed straight down, acts on the electron. If the electron is traveling in a cardinal compass direction, what are (a) the magnitude and (b) the direction of the electron's velocity at that instant?

(a) ________ m/s

(b) i. North
ii. South
iii. East
iv. West

7.4 A boy shuffles across an outdoor nylon carpet, acquiring a positive static charge, and then fires a BB gun in the direction of the setting sun. The BB travels horizontally at a speed of 105 m/s, carrying away a charge of $+3.43\ \mu$C. If the Earth's local magnetic field is perpendicular to the velocity of the BB, and it exerts a force of 1.80×10^{-8} N on it, what is the magnitude of the magnetic field at his location?

________ T

7.5 An electrically charged bullet of mass 0.010 kg leaves the barrel of a gun traveling parallel to the ground and perpendicular to the Earth's magnetic field at 1.10e+3 m/s. In theory, what charge could the bullet carry in order to remain at a constant height? Assume that the Earth is flat, and that the magnetic field (0.500 gauss) is horizontal and points left if you are looking in the direction of the bullet's motion.

________ C

7.6 An F-15 fighter aircraft is flying at 850 m/s in a direction that makes a 30.6° angle with the Earth's magnetic field, which has a strength of 5.03×10^{-5} T. The plane has acquired a static charge of 343 μC. What is the strength of the force exerted by the field on the aircraft?

________ N

7.7 A youth shuffles across an outdoor nylon carpet, acquiring a positive static charge, and then fires a BB gun due east. The BB travels horizontally at a speed of 98.4 m/s, carrying away a charge of +8.66 μC. If the Earth's local magnetic field has a strength of 5.04×10^{-5} T, is purely horizontal, and exerts a force of 4.10×10^{-8} N on the BB, what is the declination of the field at his location? Express your answer as a small positive angle between due north and the direction of the magnetic field.

________ °

7.8 A 5.3×10^{-9} kg particle carrying a charge of 3.2×10^{-6} C is accelerated by a potential difference of 3,200 V from rest. It then passes into a uniform magnetic field of strength 4.5×10^{-3} T. (a) What is the largest force it can experience? (b) What is the smallest force it can experience?

(a) ________ N

(b) ________ N

7.9 At a certain location on the Earth's equator, the magnetic field is 3.00×10^{-5} T directed northward, and there is an electric field of 108 N/C directed downward. The acceleration of gravity at this location, 9.76 m/s^2, is slightly less than the standard value of g. A proton is moving due east, a small distance above the ground. If the charge of the proton is 1.60×10^{-19} C, and its mass is 1.67×10^{-27} kg, how fast does it need to be going so that it maintains a constant distance above the ground? Ignore the curvature of the Earth.

________ m/s

Section 9 - Interactive problem: charged particle moving in a B field

9.1 Use the information given in the interactive problem in this section to answer the following questions. Test your answers using the simulation. (a) How many angles between 0° and 180°, inclusive, will result in a force of 0 N? (b) What angle in that range will give a force equal to 0 N? (If there is more than one, give the largest.) (c) What angle between 0° and 180°, inclusive, will maximize the resulting force?

(a) i. 0
ii. 1
iii. 2
iv. 3
v. 4

(b) ________ °

(c) ________ °

Section 10 - Determining the strength of a magnetic field

10.1 A proton is moving at 2.00×10^5 m/s and begins to circle at a radius of 5.00e−3 m. What is the strength of the perpendicular magnetic field acting on the proton?

________ T

10.2 A proton travels with a speed of 8.34×10^5 m/s at an angle of 32.3° with the direction of a magnetic field. The field exerts a force of 3.00×10^{-17} N on the proton. What is the strength of the field?

________ T

10.3 A charged body has a charge of 3.33 μC. It is moving with a speed of 8320 m/s at an angle of 42.0° with a uniform magnetic field. The field exerts a force of 0.0117 N on the body. What is the strength of the magnetic field?

________ T

10.4 A television screen is illuminated using a beam of electrons traveling at 7.50×10^6 m/s in a 1.75×10^{-2} T magnetic field. If an electron is moving in the positive *x* direction and the magnetic field is oriented 55.0° from the *z* axis in the first quadrant of the *xz* plane, what are the (a) magnitude and (b) direction of the acceleration of the electron?

(a) ________ m/s^2

(b) i. Positive x
ii. Positive y
iii. Positive z
iv. Negative x
v. Negative y
vi. Negative z

10.5 A uranium atom is ionized seven times, losing seven electrons in the process. Its resulting mass is 3.90×10^{-25} kg and it is given a kinetic energy of 3000 eV. If it circles in a magnetic field at a radius of 3.00 m, what is the strength of the magnetic field?

________ T

10.6 An alpha particle ($m = 6.68\times10^{-27}$ kg, $q = 3.20\times10^{-19}$ C) is moving perpendicularly to a uniform magnetic field at a speed of 7.88×10^6 m/s. Because of the force exerted on it by the magnetic field, it is accelerating at 5.67×10^{12} m/s^2. What is the strength of the magnetic field?

________ T

Section 11 - Interactive problem: B field strength and particle motion

11.1 Use the information given in the interactive problem in this section with an initial speed of 400 m/s to find the B field strength needed to achieve the desired amount of force. Test your answer using the simulation.

________ T

Section 12 - Physics at work: velocity selector

12.1 The electric field in a velocity selector is generated by two parallel plates separated by 5.00 cm with a potential difference of 90.0 V. If the magnetic field is 6.00e−2 T, at what speed will particles pass through undeflected?

________ m/s

12.2 A velocity selector is tuned to let charges with a speed of 325 m/s pass through. If the strength of the magnetic field is 0.250 T, what is the strength of the electric field?

________ N/C

12.3 The electric field strength of a velocity selector is 345 N/C. The magnetic field strength is 1.34 T. Charged alpha particles having a mass of 6.68×10^{-27} kg pass through the selector. What is the final kinetic energy of each one?

__________ J

12.4 The oppositely charged plates of a velocity selector are separated by 4.5 cm. The uniform magnetic field is supplied by a current-carrying solenoid that encloses the selector, and it has a strength of 0.35 T. What potential difference needs to be applied across the charged plates in order to select a particle velocity of 3.2×10^{4} m/s?

__________ V

12.5 A velocity selector of length 0.25 m has a magnetic field of 5.0×10^{-3} T and an electric field of 50 N/C. Electrons are injected into the apparatus and must pass through a slit at the far end which is exactly 1.0 cm across. (a) What is the velocity of a particle that goes through the velocity selector undeflected? (b) By how much can the velocity be increased or decreased from the undeflected value and still make it through the velocity selector? Assume the following: (1) The electrons begin their journey aligned with the center of the slit. (2) The magnetic force always acts in the same direction (that is, it causes an acceleration that we can approximate as linear). This is justified so long as the slit width is much smaller than the length of the velocity selector. (3) The ratio of the ideally selected velocity and the difference between the maximum and minimum velocities is very small. This will allow you to simplify some of your calculations. The final answer will demonstrate that this assumption is also justified.

(a) __________ m/s
(b) __________ m/s

Section 13 - Circular motion of particles in magnetic fields

13.1 A proton moves perpendicularly to a uniform magnetic field of strength 0.0585 T at a speed of 15,400 m/s. What is the radius of the circle it describes?

__________ m

13.2 A charged particle moves in a concentric circular orbit inside a solenoid made of very thin wires, with a diameter of 10.4 cm, just missing the inner surface of the solenoid coils. The charge on the particle is 3.00 μC, its speed is 550 m/s, and the strength of the field inside the solenoid is 0.734 T. What is the mass of the particle?

__________ kg

13.3 How fast must an alpha particle ($m = 6.68\times10^{-27}$ kg, $q = 3.20\times10^{-19}$ C) travel perpendicularly to the Earth's magnetic field in a laboratory ($B = 5.02\times10^{-5}$ T) for it to describe a circular path of radius 2.34 cm?

__________ m/s

13.4 A mad scientist has designed a new line of highly stylish, Italian-cut Kevlar body armor suits. In order to test one of these suits he wishes to shoot himself, using a .22 caliber rifle that fires rounds weighing 3.56×10^{-3} kg with a velocity of 762 m/s. To shoot himself, he will fire the gun into a vertically oriented magnetic field of strength 1.25 T, hoping that the bullet will follow a semi-circular path, emerge from the field, and fly straight toward him, striking the body armor. In order for the bullet to hit him he needs to have the bullet circle through a small radius of 2.00 cm. Finally, he knows that in order to be deflected by the magnetic field, the flying bullet needs to be electrically charged, so the firearm has an insulating rubber grip and is wired to a static electricity generator. What magnitude of charge must be placed on the bullet for the scientist's mad scheme to succeed? (The answer will be impractically large, because the scientist is truly mad.)

__________ C

13.5 What is the period of the orbit for a particle of mass 7.00×10^{-6} kg and charge 2.00×10^{-6} C circulating in a 6.00e−2 T magnetic field?

__________ s

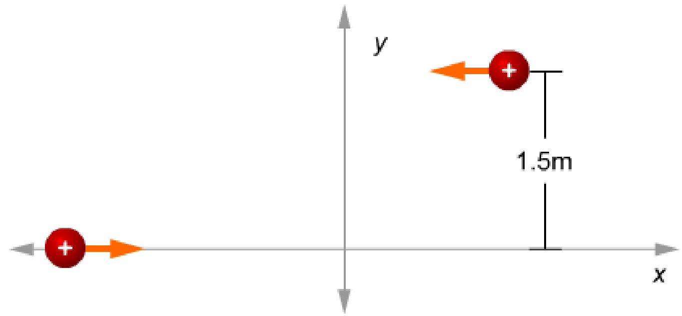
13.6 A particle with mass 1.65×10^{-12} kg and electric charge -8.22×10^{-8} C moves perpendicularly to a uniform magnetic field of strength 5.45×10^{-3} T. The angular momentum of the particle about the center of its circular orbit is 3.20×10^{-10} kg·m^2/s. (a) What is the radius of the circular orbit? (b) What is the speed of the particle?

(a) __________ m
(b) __________ m/s

13.7 An ionized hydrogen atom (that is, a free proton) is moving through the magnetic field of the Milky Way galaxy, at a location distant from any star. The kinetic energy of the proton is 5.60 keV, and it moves through a circular path having a radius of 0.0317 astronomical units. An astronomical unit (AU) equals the average radius of the Earth's orbit around the Sun, 1.50×10^{11} m. What is the strength of the galactic magnetic field in this region of space?

__________ T

13.8 A sodium ion with mass 2.32×10^{-26} kg and charge $+1.60\times10^{-19}$ C is moving with velocity 3.00×10^{5} m/s in the positive x direction. An identical particle is moving with the same speed in the negative x direction, but displaced by +1.50 m in the y direction. When the x coordinates of the atoms are equal, a magnetic field perpendicular to their motion is turned on so that the atoms collide head-on. What are (a) the magnitude and (b) the direction of the magnetic field?



(a) ________ T

(b)
i. Positive x
ii. Negative x
iii. Positive y
iv. Negative y
v. Positive z
vi. Negative z

Section 16 - Physics at work: mass spectrometer

16.1 Two elements whose atoms have the same mass number (the number of protons plus neutrons) are called isobars. The only two stable isobars with mass number 58 are ^{58}Fe with a mass of 57.933277 u and ^{58}Ni with a mass of 57.935346 u ($1\ u = 1.6605402\times10^{-27}$ kg). The mass spectrometer in a laboratory can resolve spatial displacements of 1.00×10^{-4} m after a particle's 180° flight. (a) If the particles are singly ionized and enter the spectrometer at 3.00×10^{4} m/s, what magnetic field strength will just allow these two isobars to be resolved? (Hint: for this problem, you must keep all of the significant figures for the masses of the two types of atom as you work through the problem and round your answer at the end to the correct number of significant figures. Why? Try rounding at the beginning of your work and see what happens.) (b) If you want to increase the final distance between the isobars, should you increase or decrease the magnetic field strength?

(a) ________ T

(b) ○ Increase ○ Decrease

16.2 A mass spectrometer has an adjustable magnetic field. When the magnetic field is 3.00e−2 T, a singly ionized hydrogen atom (a proton) will strike the detector at a particular location. Suppose a triply-ionized carbon 12 atom, whose mass is by definition 12 u (where $1\ u = 1.66\times10^{-27}$ kg) enters the spectrometer at the same speed as the proton. What is the strength of the magnetic field that will cause it to strike the detector at the same location as the proton?

________ T

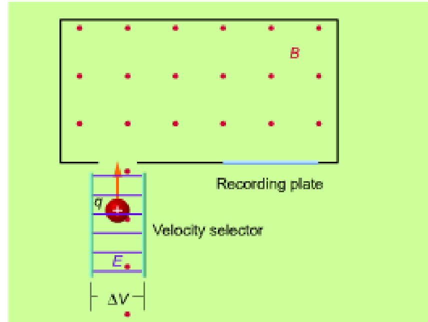
16.3 A neon atom of mass 3.82×10^{-26} kg is triply ionized. That is, three electrons are stripped from the atom so that its charge is $+3e$, or 4.80×10^{-19} C. The atom is accelerated across a potential difference of magnitude 212 V, and enters a mass spectrometer whose magnetic field has a strength of 1.25 T. (a) What is the radius of the atom's trajectory in the magnetic field? (b) What is the radius of the trajectory of a sulfur ion, also triply ionized, but having mass 5.64×10^{-26} kg? (c) What is the distance between the points on the detector plate struck by these two atoms?

(a) ________ m

(b) ________ m

(c) ________ m

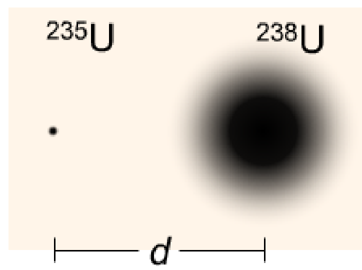
16.4 When the ionized particles entering a mass spectrometer are simply accelerated across the potential difference between two charged plates, there is always the chance of an uncontrolled variation in their velocities due to their random thermal motion **before** they undergo linear acceleration. An elaboration of the basic mass spectrometer design better controls the incoming velocity of the ions by passing them through a *velocity selector* before they enter the magnetic deflection chamber. Suppose that the same uniform magnetic field, having a strength 0.750 T, is used in both the velocity selector and the deflection chamber. In this chamber, singly ionized argon ions with a mass of 6.63×10^{-26} kg are to be deflected through semi-circular arcs of radius 12.0 cm. (a) If the oppositely charged plates of the velocity selector are separated by 2.00 cm, what is the required potential difference between them? (b) If you were to add a velocity selector to the diagram of the mass spectrometer, as shown in the illustration, which plate would be positively charged with respect to the other?



(a) __________ V

(b) ○ The right plate ○ The left plate

16.5 One of the steps in the manufacture of nuclear weapons is the separation of the fissile isotope of uranium, ^{235}U (235.05 u), from its cousin ^{238}U (238.05 u). The purity of samples of separated uranium is regularly checked in a nuclear weapons lab by vaporizing and singly ionizing the uranium, accelerating the ions and passing them through a velocity selector into a mass spectrometer deflection chamber. Photographic film is placed in the position of the recording plate. The two isotopes of uranium will strike the film plate at different locations, and their relative abundance will be indicated by the size and darkness of the two spots created on the film. The velocity selector is set to pass ions with a speed of 3.43×10^{5} m/s, and the magnetic field in the deflection chamber is 0.750 T. What is the distance *d* between the locations where the two ions of uranium strike the photographic plate? (Note that 1 u = 1.66×10^{-27} kg.)



__________ m

Section 18 - Interactive problem: mass spectrometer analysis

18.1 Use the information given in the interactive problem in this section to determine the unknown substance. Test your answer using the simulation.

i. Helium
ii. Oxygen
iii. Xenon
iv. Radon

Section 19 - Helical particle motion in magnetic fields

19.1 A particle of charge $q = 2.73\times10^{-12}$ C and mass $m = 3.32\times10^{-17}$ kg enters a uniform magnetic field **B** of magnitude 9.41×10^{-5} T with a velocity **v** of magnitude 8.34 m/s, and moves in a helix while in the field. The angle between **v** and **B** is 84.0°. (a) What is the magnitude v_{perp} of the component of **v** perpendicular to **B**? (b) What is the radius of the helix? (c) What is the magnitude v_{par} of the component of **v** parallel to **B**? (d) What is the pitch of the helix, the distance between adjacent loops (measured parallel to **B**)?

(a) __________ m/s
(b) __________ m
(c) __________ m/s
(d) __________ m

Section 21 - Interactive problem: helical particle motion

21.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) If the velocity vector is initially perpendicular to the magnetic field lines, what is the shape of the resulting path of the particle? (b) If the velocity vector is initially parallel to the magnetic field lines, what is the shape of the resulting path of the particle?

(a)
i. A helix
ii. A zigzag
iii. A circle
iv. A straight line

(b)
i. A helix
ii. A zigzag
iii. A circle
iv. A straight line

Section 22 - Magnetic force on a current-carrying wire

22.1 A 15.0 ampere current flows through a 3.75 m section of wire, perpendicular to a 0.250 T magnetic field. What is the magnitude of the force exerted by the field on the wire section?

__________ N

22.2 A straight current-carrying wire carries a current of 0.250 A. It makes an angle of 30° with a uniform magnetic field of strength 0.0255 T. What is the magnitude of the force per unit length that the field exerts on the wire?

__________ N/m

22.3 A segment of insulated wire 0.032 m long carries a current perpendicular to a uniform magnetic field. The strength of the field is 0.0055 T. The wire is attached to a spring that lies in a direction perpendicular to both the wire and the field. When a current is passed through the wire, it stretches the spring by an additional amount $\Delta x = 4.6\times10^{-4}$ m. The spring constant is 0.87 N/m. What current is passing through the wire?

__________ A

22.4 A long horizontal wire with linear mass density $\lambda = 47.5\times10^{-3}$ kg/m carries a current of 0.534 A. The wire is situated so that it is perpendicular to a uniform, horizontal magnetic field that is just strong enough to levitate the wire against its weight. What is the strength of the field?

__________ T

22.5 A wire strung across a rotating table carries a current of 0.250 A, and a uniform magnetic field of 0.500 T is applied in the plane of the table. Consider a straight section of the wire 75.0 cm long. As the table is rotated, the magnetic force exerted on this section will vary. What angle does the wire segment make with the field at the instant when the force is 0.0540 N? Report the angle that is less than 90 degrees.

__________ °

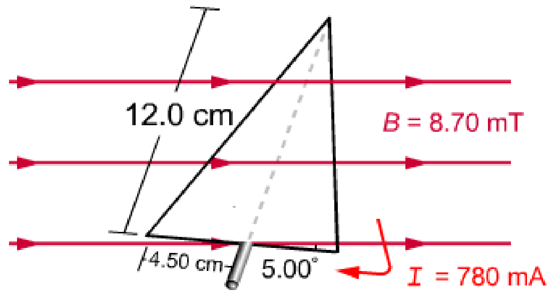
Section 24 - Torque on a loop of current

24.1 A circular wire loop of radius 0.100 m carries a current of 9.00 mA, and is placed in a 50.0 mT magnetic field. As the orientation of the loop is changed, the magnetic field exerts a varying torque on the wire loop. What is the maximum magnitude of this torque?

__________ N · m

24.2 A loop of wire in the shape of an isosceles triangle is placed in a uniform magnetic field as shown in the illustration. What is the magnitude of the torque on the loop?

__________ N · m



24.3 A circular loop of wire with a radius of 27.0 cm carries a current of 250 mA. The loop is free to rotate on a horizontal axis that is perpendicular to a horizontal magnetic field of strength 1.23 T. The loop is tilted so that its area vector makes an angle of 35.0° with the field. There is a torque on the loop which is balanced by suspending a lead weight from the loop at a point as far as possible from the axis, so that the loop does not rotate at all. What is the mass of the lead weight?

__________ kg

Section 25 - Magnetic dipole moment

25.1 What is the magnitude of the magnetic dipole moment of a circular coil of wire with a radius of 14 cm, 12 loops, and carrying a current of 340 mA?

__________ $A \cdot m^2$

25.2 A six-loop coil of current-carrying wire spanning an area of 0.0249 m^2 has a magnetic dipole moment of 6.35×10^{-3} $A{\cdot}m^2$. What current is flowing through the wire?

__________ A

25.3 Two companies manufacture magnetic dipoles made of a single wire loop carrying a fixed current. One company uses round current loops with a radius of 2.00×10^{-2} m which carry a current of 0.100 A. The other company uses square loops which carry a current of 0.150 A. What does the length of one side of the square loop have to be in order for the loops from both companies to have the same magnetic dipole moment?

__________ m

25.4 You are hiking through the Joshua Tree National Monument in California when you find to your consternation that a small grain of sand has become lodged in your compass. The pesky particle is right at the end of the compass needle, 2.15 cm from the needle's pivot point. Using a friend's compass to establish magnetic north, you hold your compass with its needle perpendicular to this direction, reasoning that the Earth's magnetic field of 5.00×10^{-5} T will apply maximum torque to it in this orientation. You tap on the compass, hoping to free up the needle. What force is the tip of the needle exerting on the grain of sand? Your compass needle has a magnetic dipole moment of 0.0188 J/T.

__________ N

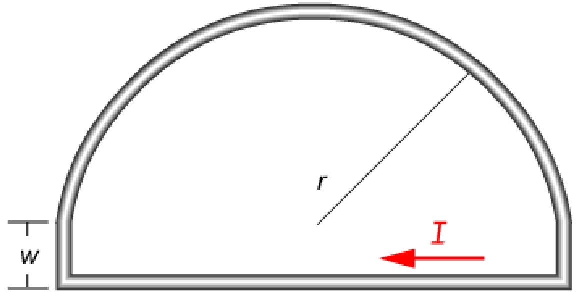
25.5 The magnetic dipole $\boldsymbol{\mu}$ has magnitude 0.125 J/T and it makes an angle θ with a uniform magnetic field **B** of strength 0.0340 T. When θ = 0°, the dipole is aligned with the field, when θ = 90°, the dipole is perpendicular to the field, and when θ = 180°, the dipole points directly opposite to the field. Calculate the magnitude of the torque exerted by the field on the dipole, and the magnetic potential energy PE_B of the dipole, for (a) θ = 60° and (b) θ = 140°.

(a) τ = __________ N·m, PE_B = __________ J
(b) τ = __________ N·m, PE_B = __________ J

25.6 As in the previous problem, the magnetic dipole $\boldsymbol{\mu}$ has magnitude 0.125 J/T and it makes an angle θ with a uniform magnetic field **B** of strength 0.0340 T. When θ = 0, the dipole is aligned with the field, when θ = 90°, the dipole is perpendicular to the field, and when θ = 180°, the dipole points directly opposite to the field. Draw graphs of the magnitude of the torque τ exerted on the dipole by the field, and the value of the magnetic PE_B of the dipole in the field for values of θ from 0° to 180°. Apply what you know about the symmetry of trigonometric functions to reduce the number of calculations you need to make.

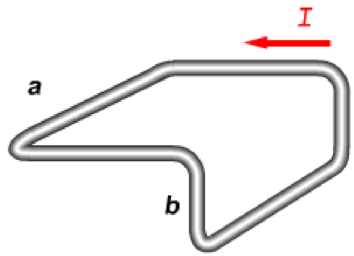
25.7 A wire is bent into the shape shown in the figure. The semi-circular part has a radius r = 4.00 cm and the rectangular part has a width w = 1.00 cm. A current of 0.300 A runs through the loop. What is the magnitude of the magnetic moment of the loop?

__________ $A \cdot m^2$



25.8 A square wire loop with sides of length a = 1.40 m is bent 90 degrees at a distance b = 0.450 m from one end, as shown in the picture. A current of 0.100 A runs through the loop. What is the magnitude of the magnetic moment of the loop?

__________ $A \cdot m^2$



Additional Problems

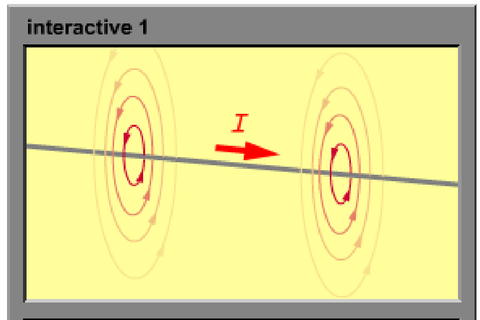
A.1 A circular wire loop of mass 2.00×10^{-3} kg and radius 1.75 cm carries a current of 0.0500 A and is rotated a small angle θ from the direction of a 3.00e−2 T magnetic field. The loop exhibits simple harmonic motion upon release. What is the angular frequency of the oscillation? (Hints: Use the small angle approximation. Also, note that the angular acceleration is directed opposite the angle. Finally, the moment of inertia of a loop of mass m about a diameter is $I = (1/2)mr^2$.)

__________ rad/s

chapter 31 Electric Currents and Magnetic Fields

31.0 - Introduction

In 1820, the Danish physicist Hans Christian Oersted (1777-1851) discovered that currents generate magnetic fields. He did so by causing a current to flow near a compass. He found that the current deflected the compass's needle; the needle aligned itself with the magnetic field produced by the current in the wire.



Vary current and observe magnetic field

His experiment proved important in both theory and in practice. It caused scientists to look for and find other relationships between electricity and magnetism. It also led to applications like electromagnets, magnets created by the flow of electricity. Such magnets are used in a host of applications, including writing data on computer hard drives.

In the simulation shown to the right, you can begin your experimentation with electricity and magnetism. You control the amount and direction of the current in a long straight wire. The magnetic field created by the current will be displayed as red-brown field lines on the screen. Arrows indicate the direction of the magnetic field. In addition, a field meter that you can move up and down, closer to and farther from the wire, displays the magnetic field strength. The vertical distance, R, between the wire and the point at which the field strength is measured is displayed above the field meter.

The simulation displays the magnetic field strength in two fashions. First, the closer together the field lines, the stronger the field. Second, the darker the field lines, the stronger the field. This approach enables you to see both how the field changes with position and how it changes with the amount of electric current.

The purpose of this simulation is for you to observe the relationship between the current and the magnetic field. You can use the simulation to determine some basic facts about currents and magnetic fields.

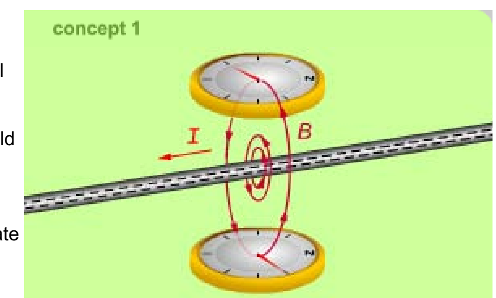
First, answer two questions about magnetic field strength: How does its strength relate to the amount of current? And to the distance from the wire?

Second, you can also study the orientation of the magnetic field. How does the orientation relate to the direction of the current? Does it change when the direction of the current changes?

To help you answer these questions and see the relationship between current and magnetic field, a controller in the simulation allows you to alter the viewing angle. You can view the wire from the side, or rotate it 90 degrees to view it end-on.

31.1 - The magnetic field around a wire

Currents produce magnetic fields. In this section, we analyze the orientation of the magnetic field created by a current flowing through a wire. In the Concept 1 and 2 diagrams, the electrons that make up the current are moving to the right. Conventional current flow is indicated (as always) with an arrow pointing in the opposite direction.

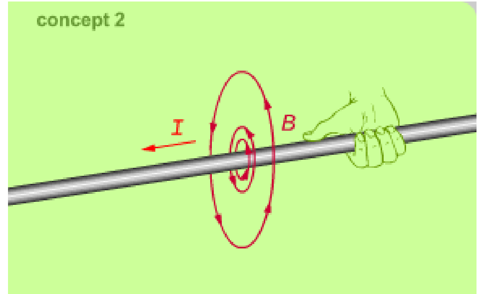


Magnetic fields and wires
Currents generate magnetic fields
· Magnetic field "circles" around wire

The Concept 1 diagram uses two compasses to show the direction of the magnetic field due to the current at two different points. Each needle points in the direction of the magnetic field at its location.

As the diagram shows, the magnetic field wraps around the current. The arrows indicate its direction. The field is strongest near the wires. As with electric field diagrams, the density of lines reflects the strength of the field. This means the lines are drawn close together near the wire.

How do you determine the direction of the magnetic field? You use the *right-hand rule for currents*. Wrap your hand around the wire so that your thumb points in the direction of the (conventional) current. Your fingers wrap around, just like the concentric circles in the diagram, and they point in the direction of the magnetic field. You see this shown in Concept 2. If the direction of the current were to reverse, your thumb would point in the opposite direction, and your fingers would wrap in the other direction. When the direction of the current changes, so does the direction of the magnetic field.



Right-hand rule for currents
Right-hand rule tells direction of field
· Point thumb in direction of conventional current
· Curled fingers show direction of magnetic field

31.2 - Two wires and their magnetic fields

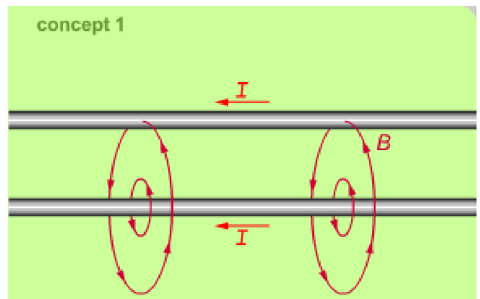
To illustrate the nature of the fields and forces produced by currents, we use the example of two wires placed parallel to one other and conducting currents flowing in the same direction, as illustrated in Concept 1. The two wires attract each other.

Let's determine why they attract each other by first considering the orientation of the magnetic field created by the wire on the bottom. We use the right-hand rule for currents to determine this orientation. This is shown in Concept 1.

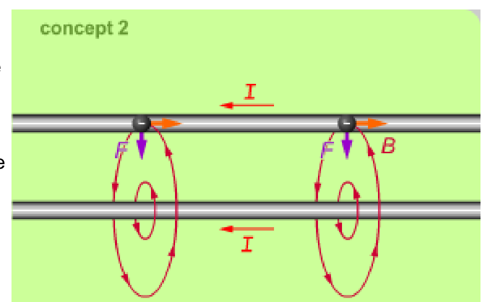
Next, we determine the direction of the force exerted by the magnetic field on the electrons moving in the upper wire. The direction of electron movement is shown with velocity vectors in Concept 2. The right-hand rule for a charge moving in a magnetic field is used to determine the direction of the magnetic force on the electron. To apply the rule, point your fingers in the direction of the velocity vector and wrap them in the direction of the magnetic field. As the diagram shows, when it intersects the upper wire, the field is pointing away from you. Wrapping your fingers from the velocity vector to the field vector will cause your thumb to point up. Since an electron is a negative charge, your thumb will point in the direction opposite to the magnetic force on it. The result is that the force points down, toward the lower wire.

The right-hand rule for force can be used again to determine the force on the moving charges in the lower wire. If you apply it, you can confirm that they experience an upward force. This analysis explains why the two wires attract each other. We show the forces acting on both wires in Concept 3.

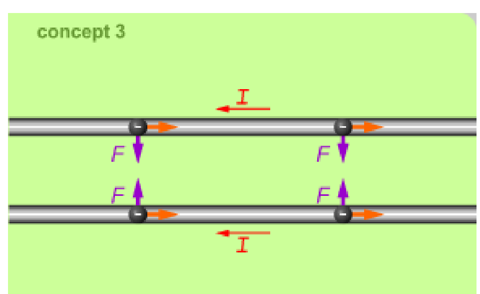
If the currents flow in opposite directions, the two wires repel each other. This scenario is illustrated in Concept 4. Here, the current in the bottom wire now flows in the opposite direction than in the prior discussion, which means its field has the opposite orientation. The force on the electrons in the upper wire now pushes them and the wire that contains them up, as the diagram reflects. The field of the upper wire pushes the electrons in the lower wire down. The two wires repel each other.



Currents and their magnetic fields
Each current creates a magnetic field
· Field intersects other wire

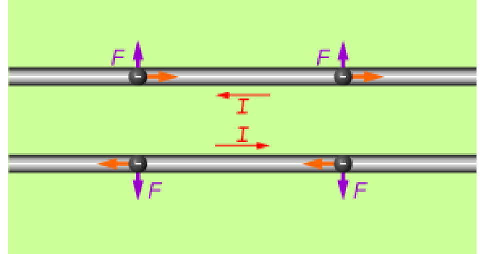


Currents in same direction
Field exerts force on moving charges
The wires attract



Currents in same direction
The wires attract

concept 4



Currents in opposite directions
The wires repel

31.3 - Strength of the magnetic field around a wire

The strength of the magnetic field around a current-carrying wire can be calculated with the equation to the right. The field strength is a function of two variables: the amount of current and the distance from the wire. The field increases with current and decreases with distance. The equation includes μ_0, the *permeability constant of free space*. This constant equals $4\pi\times10^{-7}$ N/A^2.

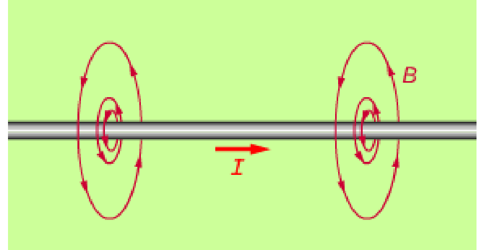
This equation is derived under the assumption that the wire is long and thin. It provides accurate values when the distance from the wire is significantly less than the wire length.

The direction of the magnetic field is determined using the right-hand rule for currents. To review: The conventional current runs to the right, so the thumb points to the right and the fingers wrap in the direction of the magnetic field.

To gain a sense of the strength of the magnetic field around a typical wire, you can look at the example problem to the right. Calculations show that the magnetic field strength 15.0 cm away from a current of two amperes is 0.0267 G.

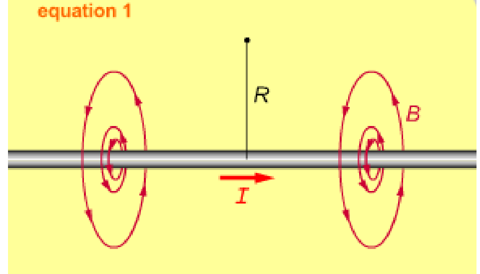
The field at this distance from the wire is much less than the average magnetic field strength of the Earth, which is about 0.5 G. On the other hand, at a point 1.5 cm from the wire, the field would be 0.267 G, or about half as strong as the Earth's field.

concept 1



Magnetic field around a wire
Strength of magnetic field
· Increases with amount of current
· Decreases with distance from wire

equation 1

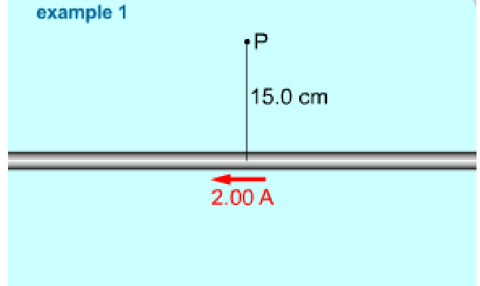


Magnetic field around a wire

$$B = \frac{\mu_0 I}{2\pi R}$$

B = magnetic field
μ_0 = permeability of free space
I = current
R = distance from center of wire
Constant $\mu_0 = 4\pi\times10^{-7}$ N/A^2

example 1



What is the magnetic field strength and direction at point P?

$$B = \frac{\mu_0 I}{2\pi R}$$

$$B = \frac{(4\pi \times 10^{-7}\ \mathrm{N/A^2})(2.00\ \mathrm{A})}{2\pi\,(0.150\ \mathrm{m})}$$

$B = 2.67\times10^{-6}\ \mathrm{T} = 2.67\times10^{-2}\ \mathrm{G}$

Field points into screen at point P

31.4 - Interactive problem: determine the current in the wire

In this simulation, there is a constant current flowing through a long, straight wire. Your task is to determine the amount and direction of this current.

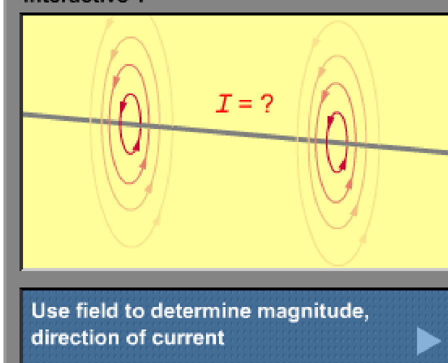
To determine the amount of current, use the magnetic field meter. It displays the strength of the magnetic field at any vertical distance R from the wire. You can drag the meter to any location to take a reading of the field strength.

To determine the direction of the current, use the right-hand rule.

Enter the amount of current, to the nearest hundredth of an ampere, in the box provided in the control panel. Select a direction for the current from the drop-down menu. Press CHECK to see whether your answers are correct.

If you have any trouble deducing the correct current, review the previous section on the strength of the magnetic field around a wire, and the earlier section that explained the right-hand rule for currents. Press RESET if you need to try again.

interactive 1



31.5 - Magnetic force, parallel wires, and the ampere

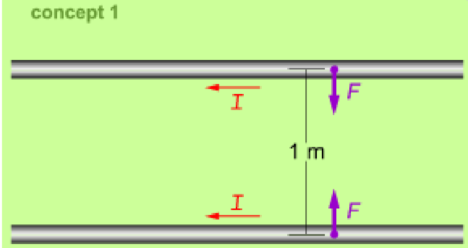
In this section, we quantify the force one current-carrying wire exerts on another, using the configuration employed by laboratories to define the ampere. Amperes are often formally defined as coulombs per second. However, to establish a practical standard for defining the ampere, scientists posit two infinitely long and infinitesimally thin, parallel straight wires, each of which carries the same amount of current. The wires are held apart at a distance of exactly one meter.

An ampere is defined as the current through the wires that creates a force per meter of 2×10^{-7} N on each wire. In the laboratory, scientists use a close physical approximation of the ideal apparatus described above.

The equation for calculating the magnetic force on each of the wires is shown in Equation 1. This force will be attractive when the directions of the currents are the same and repulsive when the currents flow in opposite directions.

We use the case of currents flowing in the same direction to derive the equation. Specifically, we will find the force the upper wire 1 exerts on the lower wire 2.

concept 1



Definition of an ampere

1 ampere in parallel wires 1 meter apart

· Creates force of 2×10^{-7} N per meter

Variables

Some of the variables used in the derivation are defined in Equation 1 on the right. Another variable used in the derivation is the strength B of the magnetic field generated by the current I_1.

The length L appearing in the equation refers to a length on **one** of the wires. F is the force exerted on an L-unit section of one current-carrying wire by the magnetic field due to the entire infinite length of the other wire.

Strategy

1. State the equation for the force exerted by the magnetic field of wire 1 on a current moving through wire 2.

2. Substitute the formula for the strength of 1's magnetic field at 2. The result is an equation that expresses the force between the wire segments in terms of their currents and geometry.

Physics principles and equations

Wire 2, and its current, are perpendicular to the magnetic field B generated by wire 1.

The force exerted by a magnetic field of strength B on a current-carrying wire perpendicular to the field is

$$F = ILB$$

where L is the length of the wire. We will also use the formula for the strength of the magnetic field generated by a current in a wire,

$$B = \frac{\mu_0 I}{2\pi R}$$

Step-by-step derivation

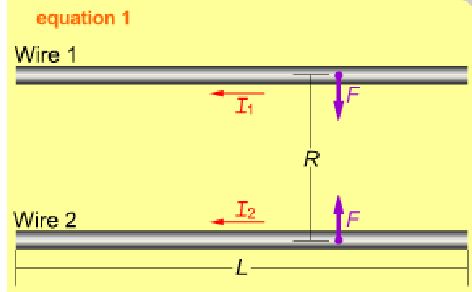
We start by stating the equation for the force on a current-carrying wire, which involves the strength *B* of wire 1's magnetic field at wire 2. Then we substitute the formula for *B* into this equation.

Step	Reason
1. $F = I_2 LB$	magnetic force on wire 2
2. $B = \frac{\mu_0 I_1}{2\pi R}$	magnetic field of a long straight wire
3. $F = \frac{\mu_0 I_1 I_2 L}{2\pi R}$	substitute equation 2 into equation 1

Even though we found the force on the lower wire, the equation derived is symmetric for the two wires. Each wire exerts the same amount of force on the other.

You can confirm the value for the force caused by one ampere of current that is stated above. If you divide the equation by the length L, you will have an expression for the force per unit length. You can then substitute one ampere for each of the currents and one meter for the distance between the wires. Using the value $4\pi \times 10^{-7}$ N/A² for the permeability constant, you will find that the force per unit length is precisely 2×10^{-7} N/m.

equation 1

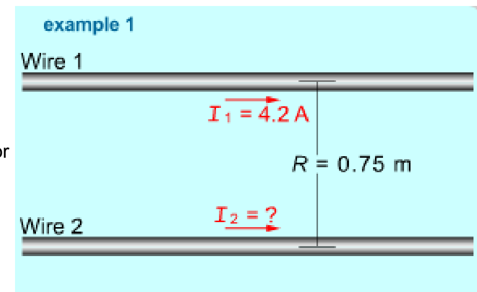


Magnetic force and parallel wires

$$F = \frac{\mu_0 I_1 I_2 L}{2\pi R}$$

F = force
μ_0 = permeability constant
I_1, I_2 = currents in wires
L = length
R = distance between wires

example 1



The force per unit length the wires exert on each other is 7.1×10⁻⁶ N/m. What is I_2?

$$F = \frac{\mu_0 I_1 I_2 L}{2\pi R}$$

$$I_2 = \frac{2\pi R}{\mu_0 I_1}\frac{F}{L}$$

$$I_2 = \frac{2\pi\,(0.75\,\mathrm{m})\,(7.1\times 10^{-6}\,\frac{\mathrm{N}}{\mathrm{m}})}{(4\pi\times 10^{-7}\,\frac{\mathrm{N}}{\mathrm{A}^2})\,(4.2\,\mathrm{A})}$$

$$I_2 = 6.3\ \mathrm{A}$$

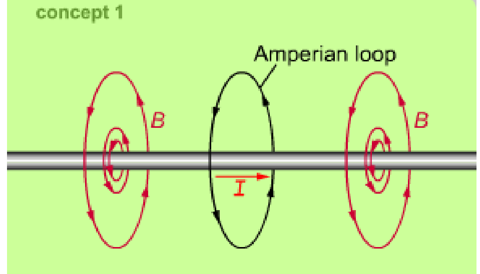
31.6 - Ampère's law

Ampère's law: The magnetic field parallel to the infinitesimal segments of an Amperian loop is proportional to the electric current passing through a surface bounded by the loop.

Ampère's law is used to determine the strength of the magnetic field caused by a current. The law uses what is called an Amperian loop, which must be a closed path: A closed path is one that starts and ends at the same point. A circle is a closed path; at the right you see a circle being used as an Amperian loop. The circle is centered on a wire that has current flowing through it.

The law states that the field strength around the loop is proportional to the enclosed

concept 1



Ampère's law
Magnetic field parallel to loop

current. The enclosed current is the current passing through a surface bounded by the Amperian loop. This method of relating current to a magnetic field was formulated by André-Marie Ampère, whose name is also used for the SI unit for current.

The equation for Ampère's law is shown on the right. It states that the component of the magnetic field parallel to the path, times the path length, equals the product of the permeability constant and the current passing through the loop. The summation sign is used because the field strength and the angle between the field and path can vary. In that case, it is necessary to choose path segments where the strength and angle are constant. The products of the parallel field component and path length for those segments can then be calculated and totaled.

The second equation shows how to calculate the component of the field parallel to the path. The cosine of θ is used to measure how much of the magnetic field lies along the direction of the path.

The equation may appear somewhat daunting, but the "trick" to applying Ampère's law is to pick an Amperian loop where the field strength and the angle are, if possible, constant. In the diagrams to the right, the loop is a circle in which the field strength is constant because all points on the circle are the same distance from the current.

The magnetic field is parallel to that circle at all points on the circle. This means the angle θ between the path and field is $0°$ at all points, and so the value of the cosine in the equation is always one. Because of this, the term on the left side of the equation reduces to the product of the constant field strength B and the whole path length s. In this case, s is the circumference of a circle (which equals $2\pi r$).

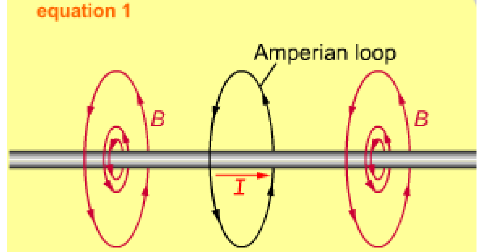
The right side of the equation contains two values that require less explanation. The first is the permeability constant, μ_0. The second is the current passing through a surface − in this case a disk is a typical choice − bounded by the path. The product of these two values equals the amount of magnetic field parallel to the Amperian loop.

The example problem on the right shows how the equation can be used to calculate the strength of the magnetic field 0.01 meters from a long, thin wire with 5.0 amperes flowing through it. The path length is the circumference of a circle with a radius of 0.01 meters.

As the solution notes, the Amperian loop is everywhere parallel to the magnetic field. The simplified equation is rearranged to solve for the magnetic field. If you study the solution, you will be able to see that the general equation stated earlier for the magnetic field caused by a long, thin wire, $B = \mu_0 I/2\pi r$, is derived as one of the steps.

proportional to enclosed current

equation 1

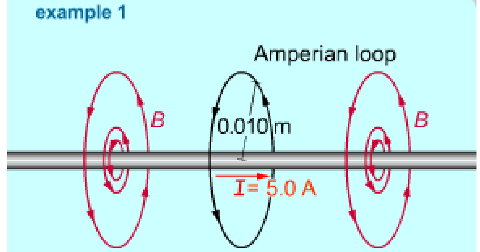


Ampère's law

$$\sum B_{\parallel} s = \mu_0 I$$
$$\sum Bs\cos\theta = \mu_0 I$$

$B_{\parallel}$ = parallel component of field
s = length of path (segment)
θ = angle between field and path
μ_0 = permeability constant
I = current
B = field strength on path

example 1



Use an Amperian loop to find the strength of the magnetic field, 0.010 m from the wire.

$\sum Bs\cos\theta = \mu_0 I$

Field is parallel to loop, so $\theta = 0°$

$Bs = \mu_0 I$

$B = \mu_0 I/s$

$s = 2\pi r$

$B = \mu_0 I/2\pi r$

$B = (4\pi\times10^{-7}\ \mathrm{N/A^2} * 5.0\ \mathrm{A})/0.020\pi\ \mathrm{m}$

$B = 1.0\times10^{-4}\ \mathrm{T}$

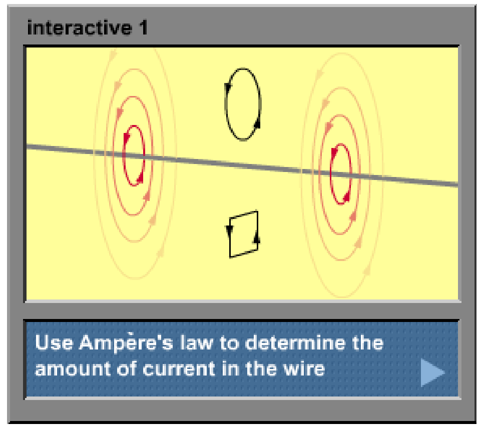
31.7 - Interactive problem: using Ampère's law

In this simulation, there is a fixed current flowing through the long, straight wire shown to the right. Your task is to use Ampère's law to determine the current.

On the screen in the simulation you will see two closed Amperian loops: one a circle, the other a square. Either closed path can be dragged up and down on the screen. At any location, a readout reports the **sum** of the components of the magnetic field parallel to the path times the length of the appropriate section of the path ($\Sigma B_{\parallel} s$).

Use either or both Amperian loops to determine the amount of current in the wire. Enter that amount, to the nearest hundredth of an ampere, in the box provided on the control panel. Press CHECK to see if your answer is correct. Press RESET to try again.

If you have trouble with this interactive problem, review the section of the textbook on Ampère's law.

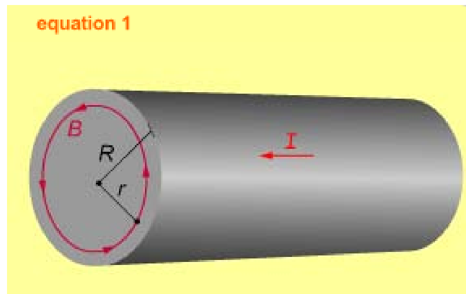


31.8 - Derivation: magnetic field inside a wire (Ampère)

In this section, we use Ampère's law to calculate the magnetic field **inside** a current-carrying wire. This wire is shown to the right. It is a straight wire assumed to have a uniformly distributed current. "Uniformly distributed" means the current density will be the same everywhere in the wire; current density is the amount of current per unit area passing through a cross section of the wire.

This derivation resembles the example worked for the field outside a wire in that it again uses a circle for the shape of the Amperian loop. As before, this causes the field strength, and the angle between the field and the path, to be constant.

Because we assume the current is uniformly distributed, we can formulate an expression describing how much current is passing through the Amperian loop as a function of the radius of the loop. As the radius increases, so does the amount of enclosed current. This is the key insight required to perform this derivation.



The magnetic field inside a wire

$$B = \frac{\mu_0 I r}{2\pi R^2}$$

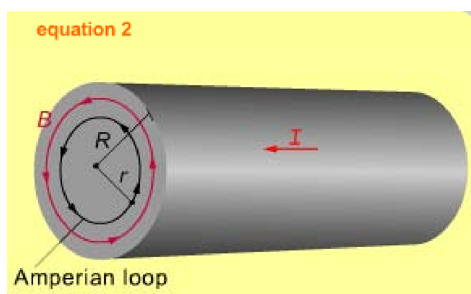
B = magnetic field
μ_0 = permeability constant
I = current
r = distance from center of wire
R = radius of wire

Variables

magnetic field strength on path	B
length of path segment	s
angle between field and path	θ
permeability constant	$\mu_0 = 4\pi\times10^{-7}$ T·m/A
current enclosed by loop	I_{enc}
current flowing through wire	I
radius of Amperian loop	r
radius of wire	R

Strategy

1. Determine the orientation of the magnetic field using the right-hand rule.
2. Based on the field, pick a shape for the Amperian loop. The path of the loop should be at a constant angle to the field.
3. Center the loop on the current so the field has a constant strength on it.
4. Use Ampère's law.
5. Take advantage of the uniform current density to relate the position of the point where the field is being measured to the amount of current passing through the Amperian loop.
6. Calculate the path length and use this value in the result from Ampère's law.



Choose an efficient Amperian loop

Physics principles and equations

The magnetic field strength will be constant at all points that are the same distance from the center of the current-carrying wire.

The magnetic field lines within the wire are concentric circles, directed as determined by the right-hand rule for currents.

These two facts, and our selection of a circular Amperian loop concentric with the wire, allow us to state Ampère's law as below:

$$\sum Bs\cos\theta = \mu_0 I$$

Uniform current density means that the amount of current passing through the Amperian loop will be proportional to its surface area.

Mathematics principles

The area of a circle of radius r is πr^2.

The circumference of a circle of radius r is $2\pi r$.

Step-by-step derivation

Having established the nature of the field and chosen an Amperian loop, we apply Ampère's law. Because we chose an efficient loop, we can evaluate the path with one expression.

Step	Reason
1. $\sum Bs\cos\theta = \mu_0 I_{enc}$	Ampère's law
2. $Bs\cos\theta = \mu_0 I_{enc}$	field strength and angle are constant along segment
3. $Bs = \mu_0 I_{enc}$	$\cos\theta = 1$

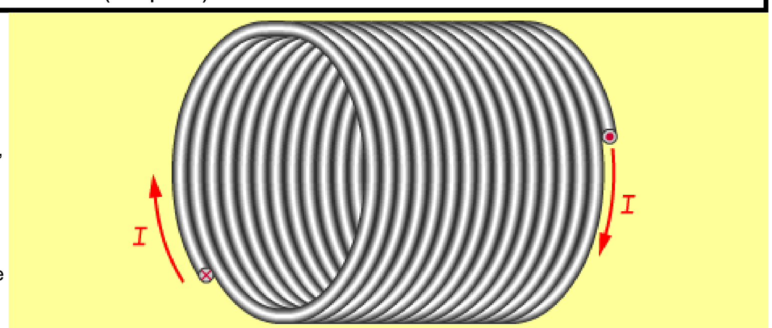
We solve the result of Ampère's law for B, using the uniform current density to find the current I_{enc} passing through the Amperian loop. We find the path length s based on the radius r of the Amperian loop.

Step	Reason
4. $B = \dfrac{\mu_0 I_{enc}}{s}$	solve for B
5. $\dfrac{I_{enc}}{I} = \dfrac{\pi r^2}{\pi R^2} = \dfrac{r^2}{R^2}$	ratio of currents equals ratio of areas
6. $I_{enc} = \dfrac{Ir^2}{R^2}$	solve for I_{enc}
7. $s = 2\pi r$	circumference of circle
8. $B = \dfrac{\mu_0 Ir}{2\pi R^2}$	substitute equations 6 and 7 into equation 4

The equation tells us that at the center of the wire − where the distance r equals zero − there is no field.

31.9 - Derivation: magnetic field of a solenoid (Ampère)

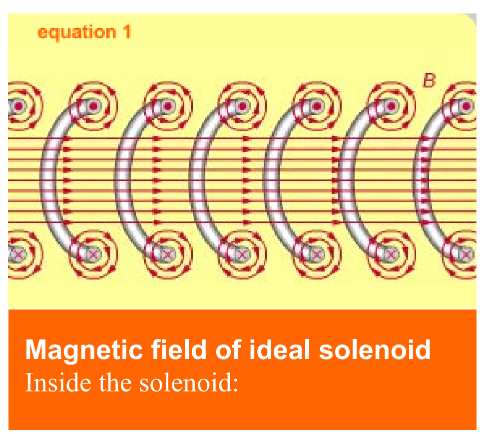
A *solenoid* consists of a wire wound into a coil of many closely packed loops of the same radius, as shown above. A solenoid, a current, and the resulting magnetic field are shown on the right. In this section, we analyze the field strength **inside** an ideal solenoid, one that is infinitely long, with tightly wrapped loops.



A section of an ideal solenoid.

The equation on the right shows the magnetic field strength inside a long, tight solenoid. The field strength depends on the current I passing through the solenoid, and the number density n of loops per unit length. We will derive this equation using Ampère's law.

Before we can derive the equation for the magnetic field, we first have to understand its nature so we can choose an efficient Amperian loop. To understand the field inside the solenoid, picture what you would see if you took many wire loops of the same size, with equal currents flowing in the same direction, and placed them side by side. This is illustrated in the first diagram to the right; here, the coils of the solenoid have been separated a bit to better exhibit the structure of its magnetic field.



Current flows through each wire loop. The red dots indicate current flowing toward you at the top, and red ×'s depict current flowing away from you at the bottom. Using the right-hand rule with any wire loop, you can verify the orientation of the circular magnetic fields shown in the illustration. Inside the solenoid, the fields reinforce each other when they are parallel to the axis (a line passing through the solenoid's center). The radial (vertical, in the cross section shown) components of adjacent fields cancel.

We will treat the magnetic field **outside** the solenoid as effectively zero when we evaluate the Amperian loop. Why is it zero? Consider a point above any loop of the solenoid. You can see why the field is zero by applying the right-hand rule to the oppositely directed currents at the top and bottom of the loop below the point. The magnetic field due to the current flowing through the top of the loop (coming toward you) is directed to the left. The field due to the current flowing through the bottom of the loop is directed oppositely, to the right.

Because one side of the loop is farther away from the point than the other, a weak field remains. However, for points far from the solenoid we use the approximation that the field outside a solenoid is zero.

To derive the magnitude of the field inside the solenoid, we enclose a portion of its upper surface with a six-segment path that has two long sides parallel to the axis of the solenoid. One of these sides is inside the solenoid, and the other is outside. This Amperian loop is shown in Equation 2.

Variables

magnetic field strength on path	B
length of n[th] path segment	s_n
angle of field to segment n	θ_n
permeability constant	$\mu_0 = 4\pi\times10^{-7}$ T·m/A
current enclosed by Amperian loop	I_{enc}
current through solenoid wire	I
wire loops inside Amperian loop	N
width of Amperian loop	L
wire loops per unit distance	n

Strategy

1. Follow the steps for determining an efficient Amperian loop. This has already been done in the discussion above, where we defined a six-segment closed path.
2. Use Ampère's law.
3. Use the definition of the *linear number density* of the loops in the coil (the number of loops per unit distance) to get the equation on the right.

Physics principles and equations

The magnetic field strength is zero outside the solenoid.

The magnetic field points to the right inside the solenoid.

These two facts, and our selection of an Amperian loop, make it convenient to use Ampère's law:

$$\sum Bs\cos\theta = \mu_0 I$$

We will also use the definition of the linear number density of wire loops

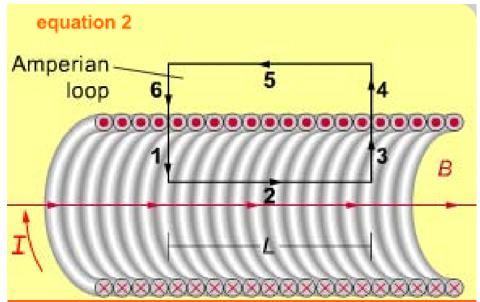
$$n = \frac{N}{L}$$

$$\boxed{B = \mu_0 I n}$$

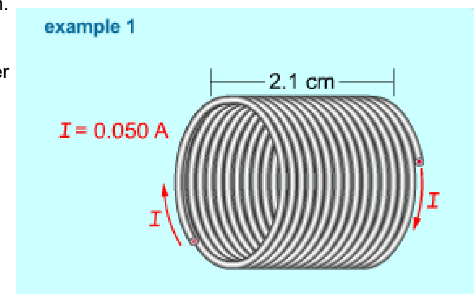
Outside the solenoid:

$$\boxed{B = 0}$$

B = magnetic field
μ_0 = permeability constant
I = current
n = wire loops per unit length



Choose an efficient Amperian loop



What is the uniform magnetic field inside this solenoid of 15 loops?

$B = \mu_0 I n$
n = number of loops / length
$n = 15 / (0.021\text{ m})$
$n = 710\text{ m}^{-1}$
$B = (4\pi\times10^{-7}\text{N/A}^2)(0.050\text{ A})(710\text{ m}^{-1})$
$B = 4.5\times10^{-5}$ T

Step-by-step derivation

Having established the nature of the field and chosen the Amperian loop, we apply Ampère's law. The loop consists of six segments, so the Amperian sum has six terms.

Step	Reason
1. $\sum Bs\cos\theta = \mu_0 I_{enc}$	Ampère's law
2. $\sum_{n=1}^{6} Bs_n \cos\theta_n = \mu_0 I_{enc}$	six path segments
3. $Bs_1 \cos\theta_1 = 0$ $Bs_3 \cos\theta_3 = 0$	$\cos\theta_1 = \cos\theta_3 = 0$
4. $Bs_4 \cos\theta_4 = 0$ $Bs_5 \cos\theta_5 = 0$ $Bs_6 \cos\theta_6 = 0$	$B = 0$
5. $Bs_2 \cos\theta_2 = BL$	$\cos\theta_2 = 1$ and segment length is L
6. $BL = \mu_0 I_{enc}$	substitute equations 3, 4 and 5 into equation 2

We replace the current I_{enc} appearing in the result of Ampère's law by the product IN, where I is the current flowing through the wire of the solenoid. Using the definition of the linear number density of wire loops, we derive the desired equation.

Step	Reason
7. $BL = \mu_0 IN$	enclosed current
8. $N = nL$	definition of linear number density of wires
9. $BL = \mu_0 I\, nL$	substitute equation 8 into equation 7
10. $B = \mu_0 I\, n$	solve for B

31.10 - Interactive problem: find the current through the solenoid

In this simulation, a 0.025-meter long solenoid is part of a circuit with a constant current, as shown to the right. Your task is to determine the amount and direction of the current.

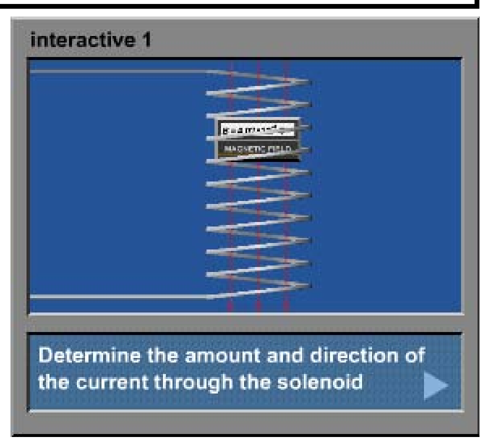
Inside the solenoid coils is a magnetic field meter. It displays the strength of the magnetic field caused by the current inside the solenoid.

The number of loops in the solenoid is displayed in a counter on the screen. You can vary the number of the loops using the up-down buttons. The length of the solenoid is fixed at 0.025 m regardless of the number of loops.

Use the magnetic field meter to determine the amount of current in the solenoid. Enter the current, to the nearest one-hundredth of an ampere, in the box provided in the control panel.

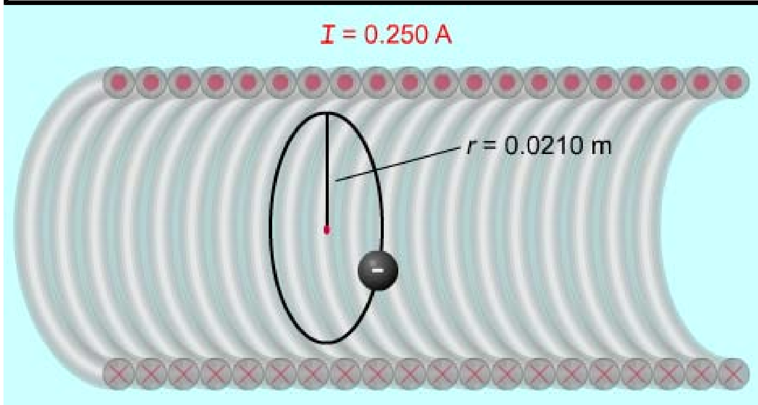
Use the right-hand rule to determine the direction of the current. Enter the direction − top to bottom or bottom to top − in the drop-down menu in the control panel.

Press CHECK to see if your answers are correct. Press RESET if you need to try again.



If you have trouble with this interactive, review the section of the textbook on the magnetic field inside a solenoid.

31.11 - Sample problem: electron in a solenoid



An electron moves in a circle, perpendicular to the magnetic field inside a solenoid. The solenoid has 1.17×10^{4} loops per meter. What is the speed of the electron?

Variables

current through solenoid	$I = 0.250$ A
loops per unit length	$n = 1.17\times10^{4}$ loops/m
radius of electron's path	$r = 0.0210$ m
electron mass	$m_e = 9.11\times10^{-31}$ kg
electron charge	$-e = -1.60\times10^{-19}$ C
magnetic field inside solenoid	B
force on electron	F
velocity of electron	v

What is the strategy?

1. The force on the moving electron is created by the magnetic field. This force is a centripetal force. Combine the equation for the magnetic force on a moving charged particle with the general equation for centripetal motion.
2. The magnetic force equation includes the strength of the magnetic field. Replace B with the equation for the magnetic field inside a solenoid.
3. Solve for the speed and evaluate the resulting expression.

Physics principles and equations

The equation for the strength of the magnetic force on a charged particle moving perpendicularly to the magnetic field is

$$F = |q|vB$$

The velocity vector of the particle is perpendicular to the magnetic field vector, since the field vector is directed along the length of the solenoid.

The equation for centripetal force is

$$F = \frac{mv^2}{r}$$

The equation for the magnetic field inside a solenoid is

$$B = \mu_0 I n$$

Step-by-step solution

Since the electron is moving perpendicularly to the magnetic field, we can use the equation for the magnetic force on a particle shown above. We combine this equation with the equation for centripetal force because the electron moves in a circle.

Step	Reason
1. $F = \lvert q \rvert vB$	magnetic force on electron
2. $F = \dfrac{mv^2}{r}$	centripetal force
3. $\lvert q \rvert vB = \dfrac{mv^2}{r}$	substitute equation 1 into equation 2
4. $B = \dfrac{mv}{\lvert q \rvert r}$	solve for B

Now we use the equation for the magnetic field inside a solenoid to relate the speed to known values.

Step	Reason
5. $B = \mu_0 In$	magnetic field inside solenoid
6. $\mu_0 In = \dfrac{mv}{\lvert q \rvert r}$	substitute equation 5 into equation 4
7. $v = \dfrac{\mu_0 In \lvert q \rvert r}{m}$	solve for speed

Finally we substitute the known values and compute the speed of the electron.

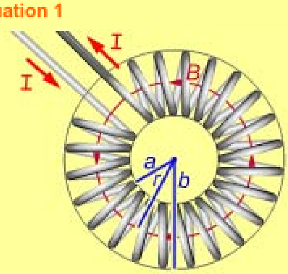
Step	Reason
8. $v = \dfrac{\mu_0 Iner}{m_e}$	electron charge and mass
9. $v = (4\pi \times 10^{-7} \frac{\text{N}}{\text{A}^2})(0.250\text{A})(1.17 \times 10^4 \frac{1}{\text{m}}) \times (1.60 \times 10^{-19}\text{C})(0.0210\text{m})/(9.11 \times 10^{-31}\text{kg})$	enter values
10. $v = 1.36 \times 10^7$ m/s	evaluate

31.12 - Derivation: magnetic field of a toroid (Ampère)

A *toroid* is a solenoid bent into a circular “donut” shape. Its magnetic field reflects its solenoid origins. A solenoid creates a uniform magnetic field that runs along its axis. A toroid is a solenoid wrapped into a circle, creating a field that circles inside the coil. The magnetic field diagram for this is shown to the right. As with an ideal solenoid, there is no magnetic field outside an ideal toroid.

Derivation. Ampère’s law provides the tools to derive the equation for the magnetic field inside the coil of the toroid. To determine the magnitude of that field, we use a circle of radius r as our Amperian loop and draw it so that it has the same center as the toroid and is enclosed by its coils. The radius r is greater than the radius a of the “donut hole” of the toroid, and less than the outer radius b of the toroid. This is illustrated in Equation 2 on the right.

equation 1



Magnetic field of a toroid

Inside the toroid loops ($a < r < b$):

$$B = \frac{\mu_0 IN}{2\pi r}$$

Outside the toroid ($r < a$, $r > b$):

Variables

magnetic field strength on path	B
length of path segment	s
angle between field and path	θ
permeability constant	$\mu_0 = 4\pi\times10^{-7}$ T·m/A
current enclosed by Amperian loop	I_{enc}
current through toroid wire	I
number of wire loops in toroid	N
radius of Amperian loop	r
inner radius of toroid	a
outer radius of toroid	b

Strategy

1. Follow the steps for determining an efficient Amperian loop. This has already been done in the discussion above, where we defined the closed circular path illustrated in Equation 2.
2. Use Ampère's law.
3. Use the configuration of the toroid and the result of Ampère's law to obtain the appropriate equation for the field strength.

Physics principles and equations

The magnetic field is oriented as shown in the diagrams to the right. The field strength is the same for all points the same distance from the center of the toroid.

These two facts, and our selection of an Amperian loop, make it convenient to use Ampère's law:

$$\sum Bs\cos\theta = \mu_0 I$$

Step-by-step derivation

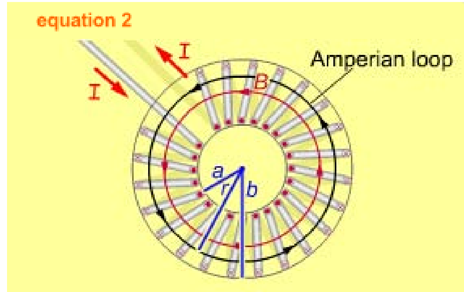
We apply Ampère's law to the selected Amperian loop. Then we make two substitutions to find the magnetic field strength in terms of the configuration of the toroid − the number of wire loops − and the radius of the Amperian loop.

Step	Reason
1. $\sum Bs\cos\theta = \mu_0 I_{\text{enc}}$	Ampère's law
2. $Bs = \mu_0 I_{\text{enc}}$	field strength and angle are constant along section
3. $s = 2\pi r$	circumference of circle
4. $I_{\text{enc}} = IN$	enclosed current
5. $B(2\pi r) = \mu_0(IN)$	substitute equations 3 and 4 into equation 2
6. $B = \dfrac{\mu_0 IN}{2\pi r}$	solve for B

An important difference between the fields of the solenoid and the toroid is that, while the field of a solenoid is uniform, the toroid's field diminishes with the distance from its center. You can see this from the equation just derived, the first formula for the strength of the magnetic field in Equation 1.

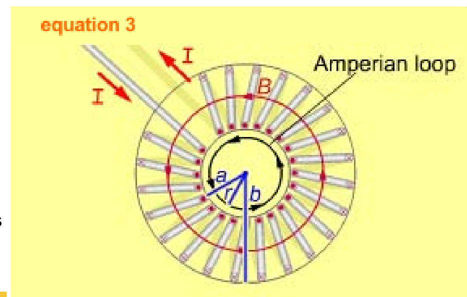
$B = 0$

B = magnetic field
μ_0 = permeability constant
I = current
N = number of wire loops
r = distance from center



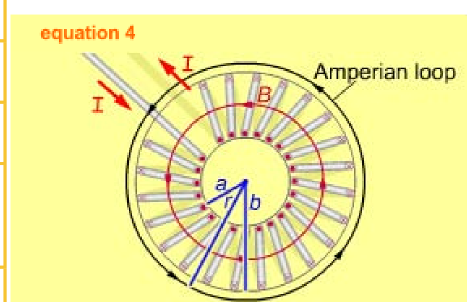
Magnetic field inside the toroid loops

Choose an Amperian loop: $a < r < b$



No magnetic field within donut hole

Choose an Amperian loop: $r < a$



No magnetic field outside the toroid

Choose an Amperian loop: $r > b$

No magnetic field within inner ring of the toroid. We can also use Ampère's law to conclude that the magnetic field inside the "donut hole" of the toroid is zero.

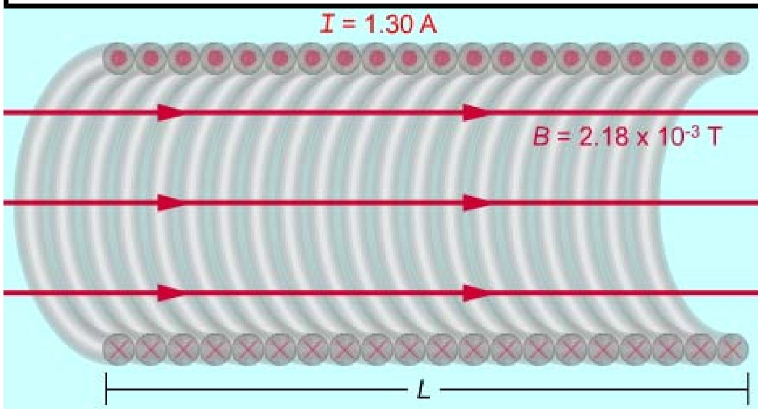
Consider an Amperian loop that is a circle inside the hole, as shown in the illustration of Equation 3. This circle is in the same midplane of the toroid as the circle used before. No current at all flows through the surface bounded by this circle, since there are no wires in the "hole." So, according to Ampère's law, there is no magnetic field along the circle sharing its orientation.

One might argue that a magnetic field could be oriented to be everywhere perpendicular to this Amperian loop, in turn making it possible for there to be a field (but no field lying along the path of the circle), and that such a field would be "invisible" to Ampère's law. To refute this scenario, we return to the right-hand rule for currents. Given the direction of the current in the inner part of each wire loop, the field produced by it would have to be in the same plane as the circle just described, not perpendicular to it. Opposing currents on opposite sides of the toroid coil do in fact produce magnetic fields parallel to the plane of the Amperian loop, **but** those fields cancel out, resulting in no net field, as we deduce from Ampère's law.

No magnetic field outside the toroid. Finally, consider an Amperian loop that is a circle outside the entire toroid, lying in its midplane, as shown in Equation 4. The wire loops of the toroid do pass through this Amperian loop, but each one passes through the circle exactly twice. There is no **net** current through the flat surface bounded by this circle. As much current flows up through the surface as down through it. Again, Ampère's law means there is no magnetic field along the circle that shares its orientation. As in the second derivation, the right-hand rule for currents and the symmetry of the toroid rule out other possible orientations of a magnetic field.

We have used Amperian loops to argue that there is no magnetic field outside the toroid coil. You could reach the same conclusion by noting that there is no magnetic field outside an ideal solenoid. It makes sense that bending a solenoid into a toroid would not cause an external field.

31.13 - Interactive checkpoint: make a solenoid



You wish to make a solenoid through which a current of 1.30 A will flow, and which will create a magnetic field of strength 2.18×10^{-3} T. If you wrap a single layer of wire as tightly as possible around a cylinder (so that there are no gaps between the loops), what thickness of wire will you need?

Consider the relationship between the number density of loops and the thickness of the wire to solve this problem. Assume that there is insulation on the wire of negligible thickness.

Answer:

d = ________ m

31.14 - Gotchas

Does Ampère's law involve the component of the field parallel or perpendicular to the path? Parallel.

A current-carrying wire passes through a surface bounded by an Amperian loop and then doubles back on itself and passes through the same surface but in the other direction. The wire creates twice as much magnetic field around the loop as it would if it had just passed through the loop one way. No, the currents combine to create no net field along the loop, since Ampère's law is based on the **net** current flowing through the surface area. The magnetic fields of the two sections of the wire are oriented in opposite directions, so along the loop, they cancel out. At a single point, the wires' fields can combine to create a net field, but along the length of the loop, the net field is zero.

31.15 - Summary

Magnetic fields exert a force on moving charges, but moving charges also create magnetic fields of their own.

A current-carrying wire creates a magnetic field that circles around the wire. The strength of the field decreases as you move farther from the wire, and increases as the current increases. The direction of the field lines is given by the right-hand rule: Your thumb points in the direction of current and your fingers wrap around the wire in the direction of the field lines.

When two current-carrying wires are placed parallel to each other, the magnetic field of each one affects the other. When the currents in the wires are in the same direction, the wires attract. When the currents are running in opposite directions, the wires repel each other.

The Biot-Savart law is used to calculate the magnetic field at a point near a given current configuration.

Ampère’s law can provide a way to calculate the magnetic field strength in situations that have some symmetry. It says that the sum of the magnetic field components around an Amperian loop (any closed path) is proportional to the net current through the loop.

It is sometimes necessary to use a line integral in Ampère’s law.

Equations

Magnetic field around a wire

$$B = \frac{\mu_0 I}{2\pi R}$$

Ampère’s law

$$\Sigma B_{\parallel} s = \mu_0 I$$

$$\Sigma Bs \cos\theta = \mu_0 I$$

Magnetic field inside a solenoid

$$B = \mu_0 In$$

Magnetic field inside a toroid

$$B = \frac{\mu_0 IN}{2\pi r}$$

Chapter 31 Problems

Chapter Assumptions

Elementary charge, $e = 1.60\times10^{-19}$ C

Mass of electron, $m_e = 9.11\times10^{-31}$ kg

Mass of proton, $m_p = 1.67\times10^{-27}$ kg

Unless stated otherwise, use 5.00×10^{-5} T for the strength of the Earth's magnetic field at its surface.

Conceptual Problems

C.1 Two parallel wires conduct equal currents in **opposite directions** (as is the case with ordinary household electric cords). At a point in the plane of the wires midway between them, (a) what is the direction of the combined magnetic field generated by the two currents? (b) Do the individual fields of the wires cancel or reinforce each other there? (c) At a point in the plane of the wires but far away to one side, do the individual fields of the wires tend to cancel or reinforce each other? Explain your answer to part c.

(a) The field is
i. parallel
ii. perpendicular
iii. at some other angle
to the plane of the wires.

(b) ○ Cancel ○ Reinforce

(c) ○ Cancel ○ Reinforce

C.2 Two parallel wires conduct equal currents in the **same direction**. (a) At a point in the plane of the wires midway between them, do the individual fields of the wires cancel or reinforce each other? (b) At a point far away from the wires, do their individual fields tend to cancel or reinforce each other?

(a) ○ Cancel ○ Reinforce

(b) ○ Cancel ○ Reinforce

C.3 Do perpendicular straight wires that carry a steady current, and are almost touching, exert a net force on each other? Explain whether the wires influence each other in any way at all.

○ Yes ○ No

C.4 Is your electric blanket safe? It has been theorized that excess nighttime exposure to electromagnetic fields may reduce the production of the hormone melatonin in certain women. Furthermore, reduced melatonin levels might lead to an increased risk of breast cancer. Scott Davis at Seattle's University of Washington School of Public Health and Community Medicine was one of the first to investigate links between electromagnetic fields and cancer (particularly leukemia). He has been quoted as saying, "Electric blankets were pretty effectively re-engineered a number of years ago. Modern-day electric blankets − anything you buy in the store now − (has a different) configuration of the heating elements in the blanket so that it doesn't produce any magnetic fields." How could the wiring in an electric blanket be designed to minimize or eliminate the production of external magnetic fields?

C.5 An up-and-coming entrepreneur, long on marketing skills but short on physics knowledge, develops a "new and improved" device for detecting current-carrying wires inside household walls. This is something a do-it-yourselfer really needs to know before picking up the old saber-saw and starting to cut out that new doorway! The "Mag-Flag Juicefinder," which costs $79.99, is actually nothing but a magnetized needle free to rotate. The advertising materials explain that the wires in the walls generate magnetic fields, and that the needle will align with these fields in a direction perpendicular to the hidden wires. Why will this device not work? Hint: Household wiring cables inside the walls contain three wires bundled close together. One is a "ground," and the same amount of current flows in opposite directions through the other two.

C.6 If you move at a constant velocity parallel to the current-carrying electrons in a straight wire and at a speed equal to their drift speed, they will not seem to be moving. In your frame of reference, is there a magnetic field around the wire? Explain your answer.

○ Yes ○ No

C.7 A Slinky® (a popular toy consisting of a helical metal coil) is hung from a high support so that each of its loops is slightly separated from the next. A strong electric current is passed through the coiled wire of this toy. Does it expand or contract? Explain your reasoning.

○ Expands ○ Contracts

C.8 A coaxial cable consists of a central current-carrying wire nested concentrically inside a conducting cylindrical sheath, which carries an equal current in the opposite direction. When the cable is in operation, is there a magnetic field outside the sheath? Explain your answer.

○ Yes ○ No

C.9 You can create a topological object called a *Mobius strip* in the following way: Cut out a long thin strip of paper, then give it a half-twist and bend it around into a loop, taping the ends together. The resulting three-dimensional figure has just one side and just one edge, as you can verify by coloring its one edge continuously with a marker; you will be able to color the whole edge without lifting the marker from the paper. The result of your coloring project is a closed loop in three dimensional space. Can the closed loop, together with the Mobius strip, form an Amperian loop and its surface?

○ Yes ○ No

C.10 The strength of the magnetic field inside a toroid varies inversely with the radius from the center point of the toroid, and everywhere outside the toroid it is essentially zero. Describe a configuration of current-carrying wires that has the opposite quality: Outside a certain toroidal region it varies inversely with the radial distance from a certain straight line, and inside the toroidal region it is zero.

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) How does the magnetic field strength relate to the amount of current? (b) How does magnetic field strength relate to the distance from the wire? (c) Does the magnetic field change direction when the direction of the current changes? (d) How does the orientation of the magnetic field relate to the direction of the current?

(a) The field strength
i. increases
ii. decreases
iii. does not change
as the current increases.

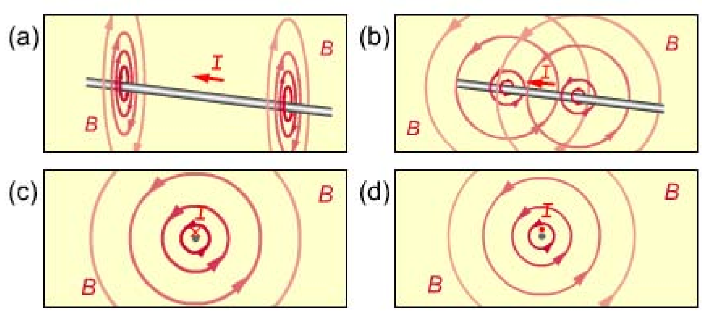
(b) The field strength
i. is linearly proportional
ii. is inversely proportional
iii. has some other relationship
to the distance.

(c) ○ Yes ○ No

(d) When the current direction is away from you, the orientation of the magnetic field is
i. clockwise
ii. counterclockwise

Section 1 - The magnetic field around a wire

1.1 The four diagrams in the illustration show a current-carrying electric wire from various viewing angles. In which diagram or diagrams is the orientation of the resulting magnetic field correctly drawn? Check all that apply.



☐ (a)
☐ (b)
☐ (c)
☐ (d)

Section 3 - Strength of the magnetic field around a wire

3.1 A long power line has a steady current of 125 A flowing through it. What is the strength of the resulting magnetic field at a distance of 1.25 m from the wire?

__________ T

3.2 A typical power transmission line carries a current of 1000 A about 30.0 meters above the ground. The magnetic field generated at ground level by the current is a matter of concern to those who live nearby. (a) What is the strength of the field at ground level below one such wire? (b) The strength of the Earth's magnetic field is about 5×10^{-5} T at sea level in middle latitudes. How close would you have to be to the power line to experience from it a magnetic field of equal strength?

(a) __________ T
(b) __________ m

3.3 When a lightning bolt strikes a lightning rod, its current is typically conducted to the ground by a ground rod. When the current in the rod exceeds 20.0 kA, it can no longer be confined by the rod and will arc laterally through the air to the ground. What is the strength of the magnetic field 3.50 m from a ground rod that is conducting current at the limit of its capabilities? Treat the ground rod as a long, straight wire.

__________ T

3.4 A continuous stream of electrons flows along a long straight path through a piece of laboratory equipment: 4.25e+18 of them pass through the equipment each second. What is the strength of the magnetic field created by this stream at a distance of 0.250 m?

__________ T

3.5 You are holding a compass 3.50 m from a long straight electrical supply wire that is carrying a direct current perpendicularly to the local direction of the Earth's magnetic field. Instead of orienting to any consistent "north" direction, the compass needle swings aimlessly no matter how you tilt or shake the instrument. The strength of the Earth's magnetic field at your location is 5.12×10^{-5} T. What is the magnitude of the current carried by the wire?

__________ A

3.6 A long, straight current-carrying wire generates a magnetic field of 1.00 μT at a distance of 1.00 m from the wire. (a) At what distance is the field strength equal to 0.100 μT? (b) At what distance is the field strength equal to 1.00 mT? (c) What current through the wire creates a 1.00 mT magnetic field at a distance of 1.00 m?

(a) __________ m
(b) __________ m
(c) __________ A

3.7 A straight horizontal wire carries a current of 3.55 μA. A proton is moving above the wire in a direction parallel to it but opposite to the direction of the conventional current in the wire, at a constant velocity of 8.00×10^{6} m/s. Assume that the force of gravity acts on the proton and that the proton is moving parallel to the Earth's magnetic field, so that the only magnetic force on it comes from the field generated by the current in the wire. (a) What is the magnitude of the upward force that the magnetic field must exert on the proton to keep it at a constant velocity? (b) How far is the proton above the wire?

(a) __________ N
(b) __________ m

3.8 A straight horizontal wire carries a current of 6.50 μA. An electron is moving above the wire a direction parallel to it and in the same direction as the conventional current in the wire, at a constant velocity of 5.41×10^{5} m/s. Assume that the force of gravity acts on the electron and that the electron is moving parallel to the Earth's magnetic field, so that the only magnetic force on it comes from the field generated by the current in the wire. (a) What is the magnitude of the upward force that the magnetic field must exert on the electron to keep it at a constant velocity? (b) How far is the electron above the wire?

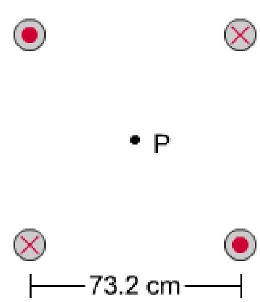
(a) __________ N
(b) __________ m

3.9 An electron is initially moving at a speed of 1.79e+3 m/s directly toward a long wire that has 2.00 A of current flowing through it. The electron is 2.50 m away. What are the direction and amount of the force exerted by the magnetic field on the electron?

__________ N, in the i. opposite ii. same direction as the current

3.10 Four parallel wires are arranged as shown in the diagram, at the vertices of a square. In two of them, a 325 mA current flows toward you, and in the other two the same current flows away from you. What is the strength of the magnetic field at the point P, equidistant from all four wires?

__________ T

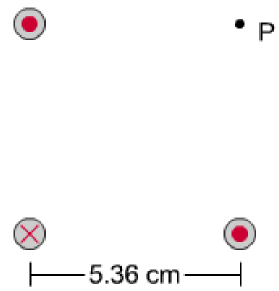


3.11 Two parallel wires are 1.50 m apart. The same amount of current will flow through each. You want a magnetic field of 4.00×10^{-5} T at the midpoint between them. (a) Must the currents flow in the same direction, or in opposite directions, through these wires? (b) How much current must flow in each wire?

(a) ○ Same direction ○ Opposite directions
(b) __________ A

3.12 Three parallel wires are arranged in a partial square configuration as shown in the diagram. In two of the wires, a 5.00×10^{-4} A current flows toward you, and in the third wire a 7.50×10^{-4} A current flows away from you. What is the strength of the magnetic field at the point P, at the "missing corner" of the square?

__________ T



3.13 A long wire is placed in a uniform magnetic field whose strength is 0.500 mT. 22.0 A of current flow through the wire, which is perpendicular to the magnetic field. At point P, the net magnetic field equals zero. How far is point P from the wire?

__________ m

3.14 Two parallel long straight wires are oriented in a north-south direction at the same height, 50.0 cm apart, and carrying currents of 8.00 A in the same direction. Find the strength of the magnetic field (a) midway between the wires, (b) at a point in the plane of the wires but 3.00 m from the nearer of the two, and (c) at a point between and above the two wires, 1.01 m from each one.

(a) __________ T
(b) __________ T
(c) __________ T

3.15 Two parallel long straight wires are oriented in a north-south direction at the same height, 10.0 m apart. The "western" wire carries 6.00 A of current to the north. The "eastern" wire carries 3.00 A of current, also to the north. There is an imaginary north-south line somewhere between the wires − not halfway! − where their magnetic fields exactly cancel out. How far is this imaginary line from the western wire?

__________ m

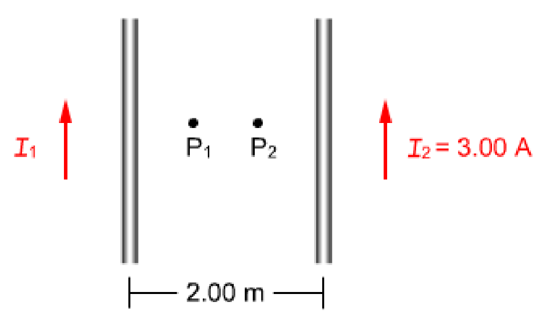
3.16 Two long straight parallel wires are oriented in a north-south direction at the same height, 50.0 cm apart, and carrying currents of 8.00 A in opposite directions. Find the strength of the magnetic field (a) midway between the wires, (b) at a point in the plane of the wires but 3.00 m to the west of the "western" wire, and (c) at a point between and above the two wires, 10.0 m from each one.

(a) __________ T
(b) __________ T
(c) __________ T

3.17 Two long, straight wires lie on the *x* and *y* axes. They conduct currents of 3.00 A and 7.00 A in the positive *x* and *y* directions respectively. What is the direction and strength of the magnetic field generated by the two currents at the point (−5.00 m, 7.00 m)?

Direction: ○ $+x$ ○ $-x$ ○ $+y$ ○ $-y$ ○ $+z$ ○ $-z$
Strength: __________ T

3.18 Two long parallel wires are separated by a distance of 2.00 m. The two points P_1 and P_2 divide that distance into three equal parts. Each wire carries a current in the same direction. The current I_2 is 3.00 A, as shown in the illustration, and the current I_1 is "tuned" so that the strength of the magnetic field at the point P_1 is zero. (a) What is the strength of the resulting magnetic field at P_2? (b) Suppose the current I_1 were re-tuned to produce zero magnetic field at P_2. What would be the resulting magnetic field at P_1?



(a) __________ T
(b) __________ T

3.19 A long straight wire lies along the *x* axis, carrying a current of 5.00 A in the positive *x* direction. Express as a three-component vector the magnetic field due to this current at the point (3.00 m, 3.00 m, 4.00 m) in three dimensional space. To orient the coordinate axes for your answer, imagine looking along the positive *x* axis, with the *y* axis straight up and the *z* axis toward the right.

(__________ , __________ , __________) T

Section 4 - Interactive problem: determine the current in the wire

4.1 Use the information in the interactive problem in this section to calculate the magnitude and direction of the current in the wire.

__________ A, i. Right to left
ii. Left to right

Section 5 - Magnetic force, parallel wires, and the ampere

5.1 Two parallel wires are separated by a distance of 5.00 cm. Each wire carries a current of 2.76 A in the same direction. (a) Find the magnitude of the force per unit length exerted by one of the wires on the other. Is the force attractive or repulsive? (b) If the currents flow in opposite directions, find the magnitude of the force per unit length exerted by one of the wires on the other. Is the force attractive or repulsive?

(a) __________ N/m, i. Attractive
ii. Repulsive
(b) __________ N/m, i. Attractive
ii. Repulsive

5.2 A weight-lifter is able to hold a barbell with a maximum total mass of 205 kg above his head. A physicist uses him to conduct a classroom demonstration, asking him to hold one insulated conducting rod (without any weights on it) above his head, while a second, very long rod is placed under his feet, parallel to the first rod. The rod that he is "lifting" has a length of 2.00 m and, together with its electrical leads, a mass of 10.0 kg. The rods are separated by a distance of 2.15 m. What is the maximum current that can be passed through the rods, in the same direction, before he is forced to yield? The same amount of current flows through both rods.

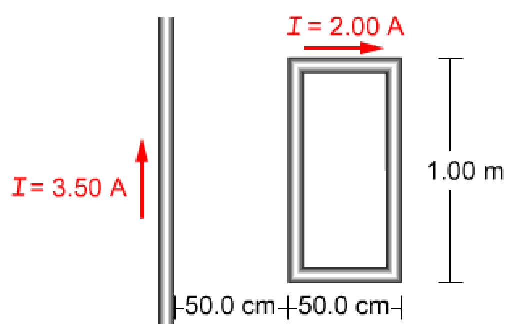
__________ A

5.3 The two conducting wires in an ordinary household electrical cord are encased in vinyl insulators. The insulated wires are connected by a thin strip of vinyl (the groove in the cord) that will burst apart if subjected to a separating force that exceeds 13,500 N/m. The conducting wires are 3.50 mm apart, the cord is stretched straight, and the same amount of direct current flows in each wire, in opposite directions. What is the maximum current, in principle, that each wire can conduct without the cord flying apart?

__________ A

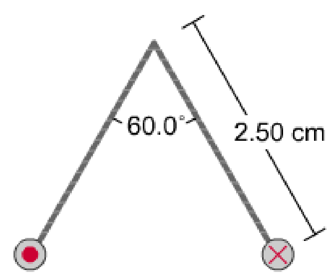
5.4 In the diagram, one current flows through the straight wire and another current flows around the wire loop, with the magnitudes shown. What are the magnitude and direction of the force exerted on the wire loop by the current in the straight wire?

__________ N, i. towards
ii. away from
the straight wire

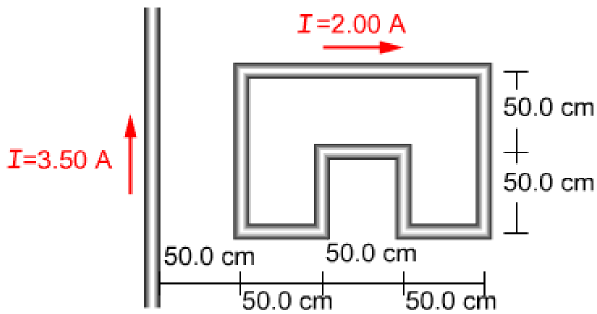


5.5 Two long straight wires with a linear mass density of 0.0850 kg/m are hung from pairs of massless strings of equal length, as shown from an end-on view in the diagram. The wires conduct equal currents in opposite directions, so they repel each other. The force of their magnetic fields opposes the tension that pulls them together. The current is adjusted until the strings reach the configuration shown and the wires are in equilibrium. What is the current?

__________ A



5.6 In the diagram, a current flows through the straight wire and another flows around the complex wire loop, with the magnitudes shown. What are the magnitude and direction of the force exerted on the wire loop by the current in the straight wire?



__________ N, i. toward
ii. away from

the straight wire

5.7 You have three parallel conducting rods. Two of them are very long, and the third is 10.0 m long, with a weight of 36.0 N. You wish to conduct the following levitation demonstration: The two long rods will be placed in a fixed horizontal orientation at the same height, 10.0 cm apart. The third rod is to float above and midway between them, 10.0 cm from each one (from an end view, the three rods will form the vertices of an equilateral triangle). You will arrange to pass the same current *I* through all three rods, in the same direction through the two "supporting" rods and in the opposite direction through the "levitating" rod. What is the magnitude *I* of the current that will maintain this astounding configuration?

__________ A

5.8 You have the same three rods as in the previous problem. After the success of your first levitation demonstration, you wish to conduct another. The two long fixed rods will still be placed in a horizontal orientation at the same height, 10.0 cm apart. But this time the third rod is to float **below** and midway between them, 20.0 cm from each one (from an end view, the three rods will form the vertices of an inverted isoceles triangle). Again you arrange to pass the same current *I* through all three rods, in the same direction through the two "supporting" rods and in the **same** direction through the "levitating" rod. (a) What is the magnitude *I* of the current that will maintain this configuration? (b) In the "repulsive force" demonstration of the previous problem, the levitating rod was resting in **stable equilibrium**: If disturbed, it would return back to its original position, like a marble resting at the bottom of a bowl. In this "attractive force" demonstration, the levitating rod is balanced in **unstable equilibrium**: If disturbed, it will move away from equilibrium, like a marble balanced on top of an inverted bowl. Why is it in unstable equilibrium?

(a) __________ A
(b)

Section 7 - Interactive problem: using Ampère’s law

7.1 Use the information given in the interactive problem in this section to determine the current through the wire. Test your answer using the simulation.

__________ A

Section 8 - Derivation: magnetic field inside a wire (Ampère)

8.1 A solid wire with a radius of 3.00 mm carries a current of 0.0530 A. Assume the current is uniformly distributed in the wire. (a) What is the magnetic field strength at the center of the wire? (b) What is the magnetic field strength at a distance of 1.00 mm from the center? (c) What is the magnetic field strength at a distance of 0.00170 m from the center?

(a) __________ T
(b) __________ T
(c) __________ T

8.2 A solid wire of radius *a* carries a uniformly distributed current *I*. (a) Use the formula for the magnetic field inside a wire to derive an equation for the strength of the field at the surface of the wire. (b) Use the formula for the magnetic field outside a wire to derive an equation for the strength of the field at the surface of the wire. (c) Are the equations the same?

(a) Submit answer on paper.
(b) Submit answer on paper.
(c) The equations i. are the same.
ii. are not

8.3 A solid conducting rod with radius 0.0100 m carries a uniform current of 2.00 A. Draw a graph of the magnetic field strength inside and outside the wire as a function of the distance *x* from the center of the wire. Use a coordinate system where the horizontal axis represents distance, and extends from 0 m to 0.0600 m, and where the vertical axis represents magnetic field strength, with an appropriate range of values in teslas.

Section 9 - Derivation: magnetic field of a solenoid (Ampère)

9.1 A long solenoid has 1.56e+3 turns in a length of 40.0 cm. What current has to flow through the solenoid in order to generate a magnetic field of 3.45 μT in its interior?

__________ A

9.2 A solenoid has a radius of 4.00 cm and a length of 50.0 cm. The number density of the windings is 300 loops per meter. A current of 0.575 A flows through the solenoid. (a) What is the strength of the magnetic field inside the solenoid? (b) What is the strength of the magnetic field outside the solenoid, at a distance of 5.00 m from its center? (c) What is the strength of the magnetic field inside the solenoid if its radius is increased to 8.00 cm, and all the other quantities have the same value as before?

(a) __________ T
(b) __________ T
(c) __________ T

9.3 A long solenoid has 2.10e+3 turns in a length of 75.0 cm. A current of 5.65 mA flows through the solenoid. What is the strength of the magnetic field generated inside the solenoid?

__________ T

9.4 A metal spring with a current passing through its loops will act like a solenoid. A current of 2.50 A passes through the (insulated) wire of the spring, and the number density of its loops when it is at its equilibirum position is 600 loops per meter. (a) What is the strength of the magnetic field generated inside this solenoid when it is at its equilibrium position? (b) What will the strength of the field be if the spring is stretched to 150% of its equilibrium length?

(a) __________ T
(b) __________ T

9.5 If one ideal solenoid is placed coaxially inside another one, is it possible to pass currents through the two solenoids so that the magnetic field inside the inner one is zero? Explain your reasoning.

○ Yes ○ No

9.6 An electron is moving in a concentric circular orbit inside a solenoid of radius 4.89 cm, in such a manner that it grazes the inside surface of the coil at all the points of its orbit. The speed of the electron is 9,820 m/s, and the solenoid has 3.50×10^3 turns/m. Assume the coils of the solenoid are so thin the electron moves essentially at the radius of the solenoid. What electric current is passing through the solenoid?

__________ A

9.7 A long solenoid with 1.20×10^3 loops per meter carries a current of 0.345 A in a clockwise direction as viewed from one end. A simple square loop with sides 3.00 cm long fits neatly inside the solenoid, its plane parallel to the loops of the solenoid. The square carries a current of 0.500 A, also in a clockwise direction. (a) What is the magnitude of the force exerted on each edge of the loop? (b) What is the magnitude of the net force on the loop? (c) What is the magnitude of the torque on the loop, considering the center of the square as the origin?

(a) __________ N
(b) __________ N
(c) __________ N · m

Section 10 - Interactive problem: find the current through the solenoid

10.1 Use the information given in the interactive problem in this section to calculate the magnitude and direction of the current flowing through the solenoid. Test your answer using the simulation.

__________ A,
i. Top to bottom
ii. Bottom to top

Section 11 - Sample problem: electron in a solenoid

11.1 A current of 0.750 A passes through a solenoid wound with 150 turns per centimeter. A proton moves around a circular path inside the solenoid, perpendicular to its magnetic field, at a speed of 2.75×10^4 m/s. What is the diameter of the circular path?

__________ m

11.2 An alpha particle (two protons and two neutrons, with a mass of 6.68×10^{-27} kg) moves in a circle inside a solenoid. The plane of the circle is perpendicular to the central axis of the solenoid, approximately parallel to the loops, and its radius is 5.00 cm. The speed of the particle is 2250 m/s. (a) What is the strength of the magnetic field inside the solenoid? (b) The solenoid carries a current of 10.0 A. How tightly wound is the solenoid, in loops per meter?

(a) __________ T
(b) __________ loops/meter

11.3 An electron is traveling inside a long solenoid in a helical path, in such a way that it exactly "tracks" the helically wound wire making up the solenoid. The solenoid has a radius of 5.25 cm and the electron stays a constant distance of 5.00 cm from the solenoid's axis. The components of the electron's velocity relative to the uniform field inside the solenoid are $v_{perpendicular} = 3.42\times10^7$ m/s and $v_{parallel}$ = 114 m/s. (a) What is the strength of the solenoid's magnetic field? (b) What is the number density of the windings of the solenoid? (c) What current must be flowing through the solenoid?

(a) __________ T
(b) __________ turns/meter
(c) __________ A

Section 12 - Derivation: magnetic field of a toroid (Ampère)

12.1 A toroid consisting of 500 loops of wire carries a current of 3.25 A. The inner and outer radii of the toroid are 2.75 cm and 4.25 cm. (a) What is the strength of the magnetic field at the inner radius of the toroid? (b) What is the strength of the magnetic field at the outer radius of the toroid? (c) What is the strength of the magnetic field at the **center** of the toroid loops? (d) What is the strength of the magnetic field at twice the outer radius of the toroid?

(a) __________ T
(b) __________ T
(c) __________ T
(d) __________ T

12.2 A toroid of 350 loops has a long straight wire passing through the center of its "doughnut hole," perpendicular to the plane of the toroid. When a current is passed through the toroid, and a larger current is passed through the straight wire in the appropriate direction, the magnetic field everywhere inside the toroid is exactly cancelled and there is a circular magnetic field at all the points outside the toroid! How many times stronger than the toroid current does the current in the long straight wire have to be?

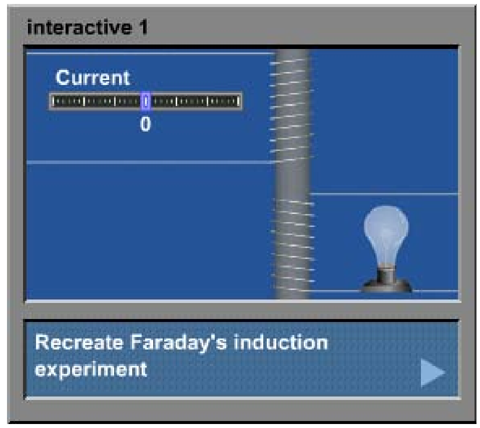
__________ times stronger

12.3 A toroid can have a square cross section as well as a circular one. Such a toroid is formed by bending a solenoid with square loops into a circle, so that the inner sides of the squares are parallel to the axis through the center of the hole of the toroid. For one particular square toroid with 350 loops, the edge of one square loop is 2.35 cm, and the inner radius of the toroid is 5.00 cm. The current through the toroid is 725 mA. (a) What is the magnetic field strength within the toroid at a radius slightly larger than the toroid's inner radius? (b) What is the magnetic field strength within the toroid at a radius slightly smaller than the toroid's outer radius?

(a) __________ T
(b) __________ T

chapter 32 Electromagnetic Induction

32.0 - Introduction



It was a chance occurrence that led the British scientist Michael Faraday (1791-1867) to discover electromagnetic induction.

It was 1831. Scientists already knew that an electric current could be used to create a magnetic field. Faraday and others were trying to achieve the opposite: They wanted to use a magnetic field to create an electric current.

Faraday was conducting an experiment in which he wrapped two lengths of insulated wire around a soft iron ring. One of the lengths was part of a circuit that included a battery. The second was part of a different circuit containing an ammeter that could measure any current passing through it. The wires were insulated so that no current could flow through the iron ring between the circuits.

Faraday knew he could create a magnetic field in the iron ring by running electricity through the first coil. His goal was to use this magnetic field to create a current in the second coil. The illustration to the right shows a modern-day recreation of his experiment. The upper coil is the part of the circuit that in Faraday's experiment contained the battery; we have replaced the battery with a slider control labeled "Current" so that you can control the amount of current in this circuit. The coil of the second circuit is the lower one in the illustration; we have included a light bulb to make it easier to see when a current flows in the second circuit.

Faraday had always connected the battery before he connected the ammeter, and he detected no current in the second circuit. However, on the morning of August 29, Faraday connected the ammeter first and then connected the battery. To his delight, he detected a momentary current in the second circuit; connecting the battery after the ammeter meant that it was measuring what happened as the current changed in the upper circuit. Today, scientists would say he created a basic transformer, and they understand that it was the **change** in the magnetic field caused by the change in current in the first circuit that caused the current in the second.

Faraday rapidly pushed his work ahead. In the next few months, he discovered that by moving a wire in a magnetic field he could also generate a current in a circuit. Today, this principle is employed in the machinery that generates most of the electricity we use.

Faraday's simple lab equipment yielded powerful insights that engineers continue to utilize today. Electric generators, microphones, VCRs, and induction stoves all rely on Faraday's discovery that a current-producing emf can be induced by changing the strength of a magnetic field, or by moving a wire in a magnetic field.

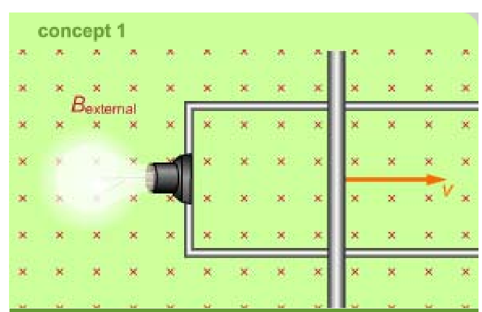
In the simulation to the right, you can recreate Michael Faraday's groundbreaking experiment. The simulation contains an experimental setup similar to the one Faraday used in 1831. You use a slider to control the current in the upper circuit on the left. By setting the slider's position, you determine the amount and direction of the current. You will see the magnetic field lines created by the current of this circuit; the more intense their color, the stronger the magnetic field.

When you launch the simulation, you will see that an oscilloscope, rather than Faraday's ammeter, is attached to the bottom circuit. It displays the potential difference across the light bulb that is part of the bottom circuit.

Experiment by changing the current in the top circuit and observing what happens in the bottom circuit. By moving the slider back and forth, you can continuously change the current. Is there a current in the bottom circuit if the current in the top circuit is steady and unchanging? What if the current in the top circuit is changing? Does the rate at which the current in the top circuit changes have any effect on the potential difference you measure in the second circuit?

32.1 - Motional electromagnetic induction

Motional electromagnetic induction: Moving a wire through a magnetic field induces an emf.



Motional electromagnetic induction
Wire moves through magnetic field
Motion induces an emf

Michael Faraday demonstrated two ways to generate a current in a circuit: by changing the strength of a magnetic field passing through a wire coil or by moving a wire through a magnetic field. This section examines the current generated by moving a wire through a field. The phenomenon is called motional electromagnetic induction, or just motional induction.

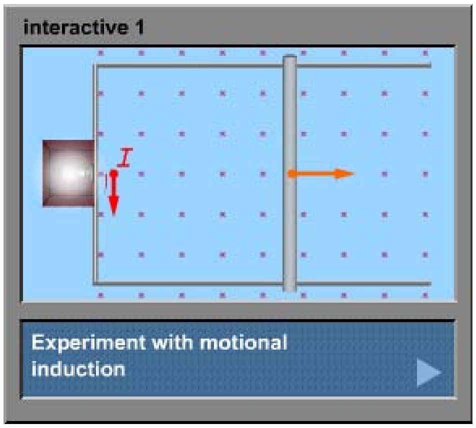
We start our explanation of motional induction by making sure the diagram on the right is clear to you. The vertical segment of conducting wire is pushed from left to right, sliding across the horizontal wires connected to the light bulb. The wires connect to form a complete circuit. The sliding vertical wire moves through an external magnetic field, represented by ×'s, that points directly into the computer screen. The sliding wire moves perpendicularly to this magnetic field.

The result is called an induced emf ($\mathcal{E}$). You studied another source of emf, a battery, earlier. The process shown here creates an emf, "induced" by sliding the vertical segment of wire through the field. Since this segment is part of a circuit, the emf causes an induced current to flow. Remember that the units of emf are volts, the same as for potential difference. We will often measure the amount of induced emf by measuring the potential difference it causes across a component like a light bulb.

Why does current flow? It flows because the motion of the wire causes the electrons in it to be moving in a magnetic field, and the magnetic field exerts a force on the moving electrons that is directed along the wire.

In the illustration on the right, you see how the motion of the sliding wire causes the bulb to glow. You may wonder: Where does the energy to illuminate the bulb come from? The answer is that it comes from whatever is doing the work of pushing the wire through the field.

32.2 - Interactive problem: motional induction



In this simulation, you can drag a wire left and right, perpendicular to a magnetic field that is pointing directly into the screen. This is the same configuration that was used to explain motional induction in the previous section. If your efforts induce an emf, current will flow and the light bulb will light.

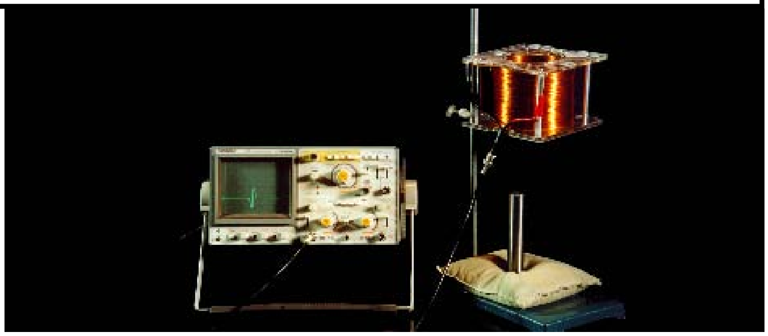
Conduct some experiments: Drag the wire slowly, and then drag it very fast. Does the speed of the wire through the magnetic field affect the amount of potential difference across the light bulb? How do the two relate? How does changing the direction you move the wire change the current? How does it change the potential difference?

You can also use the magnetic field strength control to make the magnetic field in the simulation stronger (or weaker, or oppositely directed). Does changing the field strength change the results of your efforts?

To help you to answer these questions, the simulation has an oscilloscope that measures the potential difference across the light bulb. The oscilloscope in this simulation is functionally similar to real-world ones. You can change the output scale by clicking on its dial. Initially, it is set so that one box of the display grid equals 0.5 volts, but you can change that so one box equals 0.1 volts, 10 volts, and other values shown on the dial as well. (We chose not to show the oscilloscope's connection to the circuit in this simulation for the sake of visual simplicity.) An output gauge shows the amount of current; we show its value as positive or negative to indicate direction.

There is also a slider control that lets you change the viewing angle of the simulation. This may allow you to better see the orientation of the magnetic field, and the wire moving through it.

32.3 - Induction: a coil and a magnet



Motional induction demonstration: The magnet is dropped through the coil of wire, and the resulting induced emf is displayed on the oscilloscope.

At the right, you see an apparatus often used to demonstrate induction. A bar magnet passes through a coil of wire loops and the bulb lights up.

A current flows when the magnet passes by the wire, or when the wire moves past the magnet. It does not matter which one is described as moving: The change in the magnetic field inside the coil as one moves past the other induces an emf that causes a current.

The emf induced in this demonstration will vary, based on several factors:

1. The strength of the magnetic field. The stronger the field, the greater the change in field strength as the loops move by, and the greater the induced emf.
2. The speed of the wire relative to the magnetic field. The faster one passes by the other, the greater the emf.
3. The area of the loops. The greater the area enclosed by each loop, the greater the emf.
4. The number of loops of wire. Increasing the number of loops increases the total area through which the field passes. This, too, increases the induced emf.

The list of factors above includes field strength and surface area, two of the factors that determine a quantity called magnetic field flux. We will discuss magnetic field flux shortly; it is analogous to electric field flux. Later, we will discuss how it is the rate of change of this flux that determines the amount of induced emf.

Although this is a classic way of illustrating motional induction, it is difficult to calculate the actual emf induced in this configuration. The strength and orientation of the magnetic field that intersects the coil both change as the bar magnet approaches the wire loops. It is easier to calculate the induced emf for other configurations, like a straight segment of wire moving in a uniform magnetic field.

Induced emf depends on:
Strength of magnetic field
Speed of magnetic field past wire
· Or wire past field
Area of loop
Number of loops

32.4 - Physics and music: electric guitars

The *pickups* (raised black modules) on this electric guitar convert the vibrations of the strings into a varying electric current.

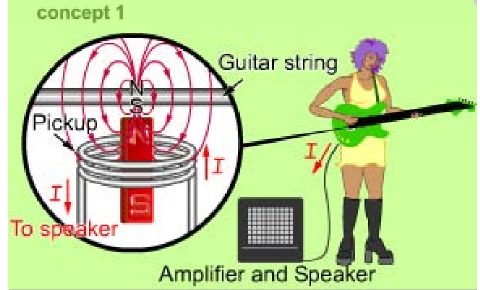
Electric guitars use electromagnetic induction in a component called a *pickup* to produce music. In the pickup, the vibrations of the strings are converted into an electrical signal so they can be amplified and then played over speakers.

The pickup consists of a permanent magnet surrounded by a coil of wire. You see a conceptual diagram of a pickup in the illustration to the right. Guitars typically have two or three pickups under every string, each pickup designed to be maximally sensitive to a particular frequency.

The permanent magnet in the pickup serves to magnetize the nearby guitar string. A musician plucks the string, making it vibrate. With this vibration the magnetized string moves back and forth near the coil, inducing an emf in the wire. This emf causes a current in the wire. The emf and the current oscillate at the same frequency as the string.

The signal is transmitted through a circuit to an amplification system, which increases its strength, or energy. The system then sends the signal to a loudspeaker and the audience hears the music.

concept 1



Physics and music: electric guitars
String magnetized by permanent magnet in pickup
Guitar player makes string vibrate
Motion of magnetized string induces current in coil
Current flows to amplifier/speaker

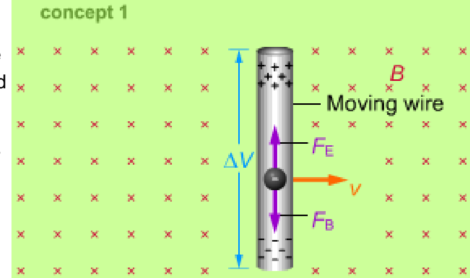
32.5 - Motional induction: calculating the potential difference

The diagram to the right shows a vertical wire segment moving through a uniform magnetic field. This field has a constant strength and points into the screen. The wire moves at a constant velocity to the right, which means its motion is perpendicular to the magnetic field. In this section, we show how to calculate the potential difference induced between the ends of the wire segment by its lateral motion.

The potential difference is caused by the force of the magnetic field on mobile electrons in the moving wire. The electrons move together with the wire as it is pushed through the field. Because charged particles experience a force when they move through a magnetic field, the mobile electrons get pushed to one end of the wire, leaving a positive charge on the other end. (Which end the electrons move to can be determined by a right-hand rule.) As more and more electrons accumulate on one end, leaving more and more positive charge on the other, there is an increasing potential difference across the wire.

The wire's motion causes the formation of two charged regions, one positive and the other negative. Now, let's consider the electron that is shown in the center of the wire in Concept 1. The motion in the magnetic field causes it to experience a downward magnetic force F_B, and the charged regions cause it to experience an upward electric force F_E. When these two forces balance, the electron experiences no net force and stays in place. The system is in equilibrium and the equation on the right can be used to determine the potential difference across the wire.

concept 1



Calculating potential difference
Potential difference induced across wire moving in a magnetic field
Electric, magnetic forces act on charge
· Forces balance

In a simulation a few sections ago you saw how, when a moving wire segment like this is part of a circuit, it becomes an emf source that can be used to illuminate a light bulb. The amount of the potential difference in this scenario equals the amount of the motional emf in the former one.

To derive an expression for the equilibrium potential difference, we start with an equation stating that the magnetic force on the charged particle equals the electric force. The other variables we use are shown in Equation 1 or defined in the strategy steps below.

Strategy

1. State that the forces are in equilibrium: $F_B = F_E$.
2. Substitute expressions for F_B and F_E, the magnetic and electric forces respectively, in terms of the quantities they depend on.
3. Use the relationship between field strength, potential difference and distance in a uniform electric field to rewrite E, and then solve for the potential difference between the ends of the wire segment.

Physics principles and equations

The strength of the force exerted by a magnetic field on a charge moving perpendicular to it is

$$F_B = qvB$$

We will also use the definition of an electric field.

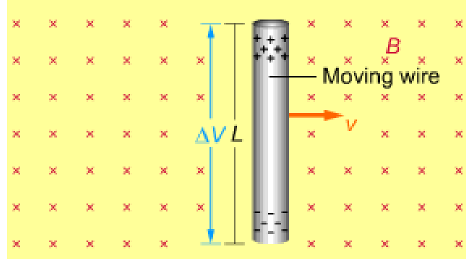
$$E = \frac{F_E}{q}$$

The equation that relates potential difference to electric field strength and distance in a uniform electric field is

$$\Delta V = Ed$$

In this case, the distance d is the wire length L.

equation 1



Potential difference across wire segment

For motion perpendicular to field:

$$\Delta V = LvB$$

ΔV = potential difference across wire
L = length of wire
v = speed
B = magnetic field strength

Step-by-step derivation

We employ the strategy and equations mentioned above and use algebra to solve for the potential difference.

Step	Reason
1. $F_B = F_E$	forces in equilibrium
2. $qvB = Eq$	substitute expressions for forces
3. $vB = E$	divide by q
4. $E = \frac{\Delta V}{L}$	solve given equation for E
5. $vB = \frac{\Delta V}{L}$	substitute equation 4 into equation 3
6. $\Delta V = LvB$	solve for ΔV

As the equation indicates, longer wires, higher velocities, and stronger magnetic fields lead to greater induced potential differences. This derivation considered the case of motion perpendicular to the magnetic field. If the motion were not perpendicular, we would use trigonometry to determine the component of the motion that was perpendicular to the field. In general, we would conclude that $\Delta V = LvB \sin \theta$, just as the magnetic force equals $qvB \sin \theta$.

32.6 - Magnetic flux

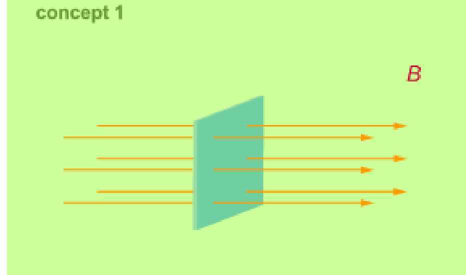
Magnetic flux is analogous to electric flux. Specifically, magnetic flux is the product of a surface area with the component of a magnetic field passing perpendicularly through the surface, just as electric flux is the measure of how much electric field passes perpendicularly through a surface. The unit for magnetic flux is the *weber* (Wb). One weber is one tesla·m^2, the units for magnetic field strength times those for area.

With both electric and magnetic flux, the cosine of the angle between the field and the area vector is used to measure the component of the field passing perpendicularly through the surface. The area vector is normal to the surface and equal in magnitude to its area.

In Equation 1, you see the equation for magnetic flux. It states that magnetic flux equals the dot product of the magnetic field and area vectors, which is calculated as the product of the magnetic field strength, the surface area, and the cosine of the angle θ. This angle is shown in the diagram.

You will use the concept of magnetic flux, and changes in magnetic flux, to further your understanding of how changing magnetic fields induce emfs.

concept 1

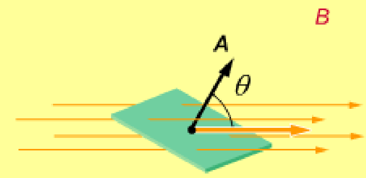


Magnetic flux

Amount of field passing through surface
Depends on:
· Field strength at surface

· Amount of surface area
· Angle between field, area vector

equation 1

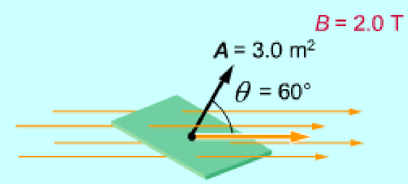


Magnetic flux

$$\Phi_B = \mathbf{B} \cdot \mathbf{A}$$
$$\Phi_B = BA \cos \theta$$

Φ_B = magnetic flux
$\mathbf{B}$ = magnetic field, $\mathbf{A}$ = area vector
θ = angle between field, area vector
Units: webers (Wb = T·m^2)

example 1



What is the magnetic flux through this surface?

$\Phi_B = BA\cos \theta$
$\Phi_B = (2.0\ \text{T})(3.0\ \text{m}^2)(\cos 60°)$
$\Phi_B = 3.0\ \text{Wb}$

32.7 - Faraday's law

Faraday discovered two ways to induce an emf. The first was by moving a wire through a magnetic field; the second was by changing the strength of the magnetic field passing through a stationary wire coil. In the latter case, the magnetic flux changed because flux is proportional to the strength of the magnetic field.

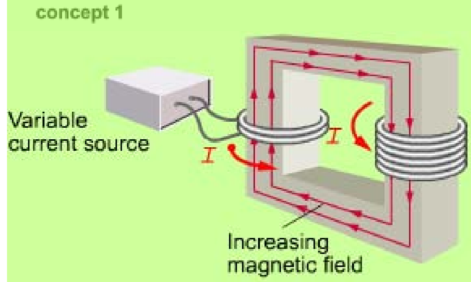
In general, since flux is the product of the magnetic field strength, the surface area, and the cosine of the angle between the field vector and the area vector, changing any of these three factors changes the flux and induces an emf.

In this chapter, we will look at all three of these ways of inducing an emf. We start by analyzing the emf induced by a change in magnetic field strength. In later sections we will consider changes in surface area and angle.

To consider changes in magnetic field strength, we use the apparatus shown to the right. It is similar to Faraday's equipment. Two wire coils are wrapped around a piece of iron. Both coils are insulated so that no current can flow directly between them. The coil on the left is connected to a variable current source, a device that can cause a current that changes over time to flow through the coil. As the current in this coil changes, so will the magnetic field that it creates.

The iron core facilitates the transmission of the magnetic field from within the left-hand coil to within the coil on the right. We use two illustrations of the same configuration to

concept 1



Changing magnetic flux and induced emf

Change in magnetic field strength yields change in magnetic flux
Flux change induces emf
emf drives induced current

show what occurs. In Concept 1, you see the overall configuration: the variable current source attached to a coil on the left, the changing magnetic field passing through the iron core, and the current that is induced in the coil on the right. In Equation 1, we show the view looking down the coil on the right. The increasing magnetic field points away from you down the coil in this view, resulting in a counterclockwise induced current.

Let's discuss in more detail what is happening in this system. The iron core ensures that the magnetic field passes essentially unchanged from within one coil to the other. Since the field is perpendicular to the loops of the coil on the right, the magnetic flux passing through this coil equals the product of the magnetic field strength, the surface area of a loop, and the number of loops in the coil.

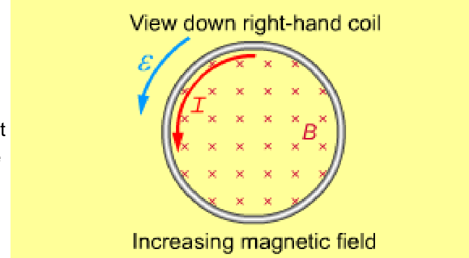
The current flowing through the coil on the left creates a magnetic field. As that current changes over time, so does the magnetic field it generates, which means the magnetic flux passing through the coil on the right changes. That change in magnetic flux induces an emf in the coil on the right, which in turn causes the current in the circuit on the right.

This process is used to illustrate a general principle, called Faraday's law: A change in magnetic flux induces an emf. In Equation 1, you see this expressed in mathematical form. Faraday's law states that the induced emf equals the negative of the rate of change of magnetic flux.

Faraday's law is often applied to a coil having N loops, as we do in this section, and we state this version in Equation 2. In this case, the flux refers to the flux passing through each loop, and the total flux equals the flux passing through each loop times the number of loops.

The negative sign appearing in both equations indicates that the induced emf acts to "oppose" the change in magnetic flux that causes it. What this means is explained in more depth in a later section.

equation 1

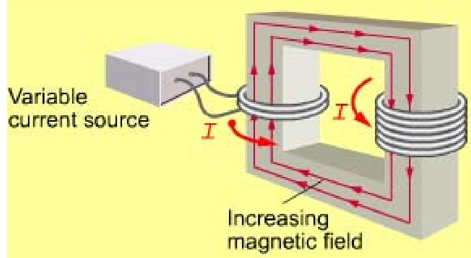


Faraday's law, total flux

$$\mathcal{E} = -\frac{\Delta\Phi_B}{\Delta t}$$

$\mathcal{E}$ = induced emf
Φ_B = total magnetic flux through circuit
t = time

equation 2

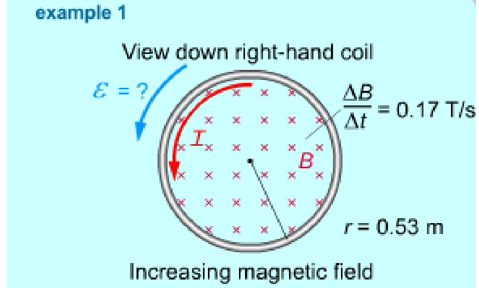


Faraday's law, coil of *N* loops

$$\mathcal{E} = -N\frac{\Delta\Phi_B}{\Delta t}$$

Φ_B = magnetic flux through one loop
N = number of loops

example 1



The field passes through six loops. What is the induced emf?

$$\mathcal{E} = -N\frac{\Delta\Phi_B}{\Delta t} = -N\frac{\Delta B}{\Delta t}A$$

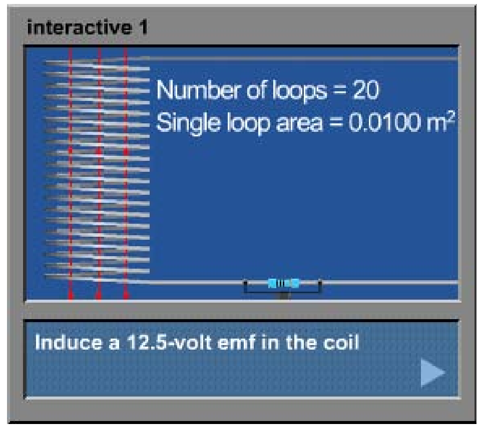
$$A = \pi r^2 = \pi(0.53\text{ m})^2 = 0.88\text{ m}^2$$

$$\frac{\Delta\Phi_B}{\Delta t} = \frac{\Delta B}{\Delta t}A = \left(0.17\frac{\text{T}}{\text{s}}\right)(0.88\text{ m}^2)$$

$$\frac{\Delta\Phi_B}{\Delta t} = 0.15\text{ V}$$

$\mathcal{E} = -(6 \text{ loops})(0.15 \text{ V}) = -0.90 \text{ V}$

32.8 - Interactive problem: Faraday's law



In this simulation, a magnetic field is passing through a solenoid. The solenoid is part of a circuit that contains a resistor. The potential difference across the resistor equals the magnitude of any emf induced in the solenoid.

You control the rate of change of the magnetic field passing through the solenoid. Your task is to use Faraday's law to calculate the time rate of change of magnetic flux through the solenoid that will induce an emf having a magnitude of 12.5 V and cause current to flow in the circuit. In turn, this will create a potential difference of 12.5 V across the resistor.

The solenoid has 20 loops, and each loop has an area of 0.0100 m^2. An oscilloscope measures the potential difference across the resistor.

The magnetic field starts at 0.500 T and will decline linearly to 0 T during a time interval you specify. You have a controller that sets the time interval during which the change in field strength will occur to values from 5.00 to 20.00 milliseconds. The field will then continue to vary back and forth between 0 T and 0.500 T during alternating time intervals of the same length. The rate of change will alternate between positive and negative values, but it will have a constant magnitude based on the duration of the time interval you select.

Specify the time during which you want the field to decline to zero. (You change it in increments of 0.10 milliseconds.) Press GO to see if the changing field induces an emf of magnitude 12.5 volts. If not, press RESET, redo your calculations and try again.

If you have trouble answering this problem, review the section on Faraday's law. The oscilloscope displays a graph of the potential difference across the resistor, with the potential difference plotted on the vertical axis and elapsed time on the horizontal axis. By clicking on different values on the oscilloscope's control knob, you can specify what you want the vertical measure, in volts, of one grid square to be.

32.9 - Flux and motional induction

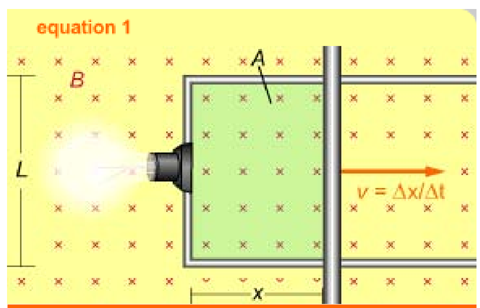
Faraday's law relates an induced emf to a rate of change of flux. In the preceding sections we discussed how a change in field strength led to a change in flux; we will now discuss how a change in surface area can lead to a change in flux and to an induced emf. In a later section we will discuss how a change in the angle between a magnetic field and a surface area vector can also induce an emf.

Earlier, we used a configuration similar to part of the one shown to the right to determine that the **potential difference** induced across a wire segment as it moves perpendicularly to a magnetic field equals LvB. Here, this segment is now part of a circuit, and we can use Faraday's law to derive the same formula for the magnitude of the **induced emf** in the segment.

You see the equation we derive in Equation 1, where we use $\mathcal{E}$ to represent the magnitude of the emf. This equals LvB, the product of the length L of the wire segment, its speed v and the strength B of the magnetic field. The wire moves perpendicularly to the field; if it moved at some other angle, we would use trigonometry to determine the component of its velocity that was perpendicular to the field.

This derivation hinges on the fact that the rate of change of the area can be defined in terms of the speed of the wire, which we assume to be constant. We use A for the area and x for the width of the rectangle enclosed by the circuit in the derivation.

The loop in this scenario is the perimeter formed by the moving segment and the other wires that form the closed circuit. Since the magnetic field passes through it perpendicularly, the flux equals the product of this area and the field strength. The area will change as the wire moves, increasing if the wire slides to the right, and decreasing if the wire slides to the left. Since the magnetic field is assumed to be uniform, the change in flux is solely due to the change in area.



Induced emf caused by moving wire

$$\mathcal{E} = LvB$$

$\mathcal{E}$ = magnitude of induced emf
L = length of wire
v = speed
B = magnetic field strength

To calculate the magnitude of the emf, we use Faraday's law without a negative sign, and the fact that the flux Φ_B equals the area of the loop A times the magnetic field B.

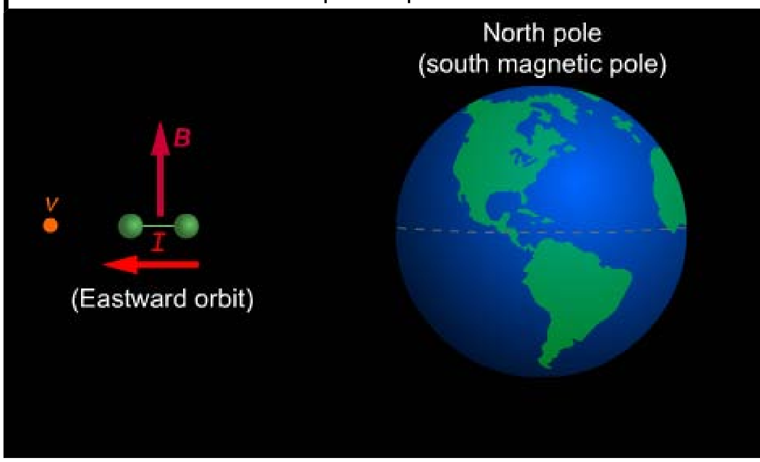
$$\mathcal{E} = \frac{\Delta \Phi_B}{\Delta t}, \qquad \Phi_B = AB$$

Magnitude of induced emf across segment

We use Faraday's law to express the magnitude of the emf in the wire. We take the rate of change of the flux to be proportional to the rate of change of the area, which is in turn proportional to the horizontal speed of the wire segment.

Step	Reason
1. $\mathcal{E} = \frac{\Delta\Phi_B}{\Delta t}$	Faraday's law
2. $\Phi_B = AB$	flux
3. $\frac{\Delta\Phi_B}{\Delta t} = \frac{\Delta A}{\Delta t} B$	divide by Δt
4. $A = Lx,\ \Delta A = L\Delta x$	definition of area
5. $\frac{\Delta A}{\Delta t} = \frac{L\Delta x}{\Delta t} = Lv$	divide by Δt; speed equals $\Delta x/\Delta t$
6. $\mathcal{E} = LvB$	combine equations

32.10 - Interactive checkpoint: space tether



As a particular tethered spacecraft orbits the Earth, its velocity (pointing out of the screen in the illustration) is perpendicular to the Earth's magnetic field, which induces an emf in the tether wire (assume that the magnetic field is locally uniform). If the emf induced between the ends of the wire is 3.50×10^3 V, and the tether is moving at 6.81×10^3 m/s perpendicular to a magnetic field of 20.0×10^{-6} T, how long is the tether?

An electrodynamic space tether, such as the one illustrated here, is a conducting cable that joins two satellites orbiting at different altitudes. Since the satellites are joined, they must orbit at the same speed. This means the inner one is "too low" for its orbital speed, and the outer one is "too high." The weight of the inner satellite exerts a tension on the tether, providing a centripetal force that keeps the outer satellite in its orbit.

Space tethers are an emerging technology that could provide low-cost electrical power and propulsion for space stations, as well as gathering unwanted "space junk." They work because a current is induced in a long conducting wire as it moves perpendicularly through the Earth's magnetic field. The induced current can power operations inside the station, but it also creates drag as the Earth's magnetic field exerts force on the current.

The current cannot flow through the space tether indefinitely, however, because the tether is not part of a circuit, and the current causes a separation of charge: The outer satellite in the illustration above will become positively charged, and the inner one will become negatively charged. As soon as the electrostatic force due to the separation of charge is strong enough to balance the magnetic force on charges in the tether, the induced current will cease flowing of its own accord.

Alternatively, by using another source of power, the station can send current through the tether in the direction opposite to the induced current. This imposed current interacts with the magnetic field to accelerate the space station and boost it into a higher orbit.

Answer:

$L =$ ____ m

32.11 - Lenz's law

Lenz's law: An induced current flows so that the magnetic field it creates opposes the change in magnetic flux that causes the current.

Lenz's law, named after the Russian scientist Heinrich Lenz (1804-1865), is used to determine the orientation of the current induced by a change in magnetic flux.

The law states that the magnetic field of the induced current opposes the change in magnetic flux that causes the current. To apply the law, you first note the change in magnetic flux and then determine the orientation of the magnetic field that will oppose that change in flux. The current will flow in the direction that causes this magnetic field.

To illustrate this law being applied, we use the wire loop shown to the right. The external magnetic field is pointing into the page and it is **increasing** in strength. This will cause a current. But in which direction does the current flow, clockwise or counterclockwise?

To answer this question, consider the direction of the magnetic field created by the induced current. Lenz's law says this field will oppose the change that caused it. The external field is increasing, so the magnetic field of the induced current points in the direction that opposes this change. This means it will point toward you, as the diagram in Concept 2 shows. (In this diagram we have dimmed the external field to make the opposing field easier to see.)

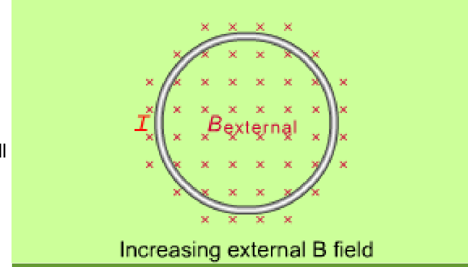
The right-hand rule for currents dictates that the current must be flowing counterclockwise through the wire loop. If you apply the rule, your fingers inside the loop must point toward you, opposing the increase of the external magnetic flux that is pointing away from you. This means the thumb points up on the right side of the loop, as illustrated in Concept 2, and down on the left side, as you can imagine, indicating the direction of conventional current. "Up on the right" and "down on the left," means the current flows counterclockwise.

(Perhaps it is useful to state a right-hand rule for induction loops: "The thumb points in the direction of the induced magnetic field, and the fingers wrap around the loop to indicate the direction of the current.")

If the external field were decreasing, then the induced magnetic field would oppose the decrease. It would point in the same direction as the external field, and this means the current would flow clockwise.

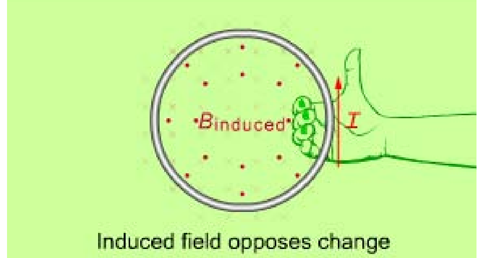
The example problem asks you to determine the direction of the current in a loop when a decreasing external magnetic field passes through it, directed toward you. Remember: The current must flow so that it creates a magnetic field inside the loop that opposes this decrease in flux.

concept 1



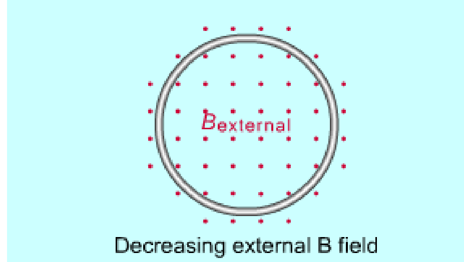
Lenz's law
Determines direction of induced current

concept 2



Lenz's law
Determines direction of induced current
Magnetic field of induced current opposes change in flux

example 1



In what direction will the current flow?
Counterclockwise

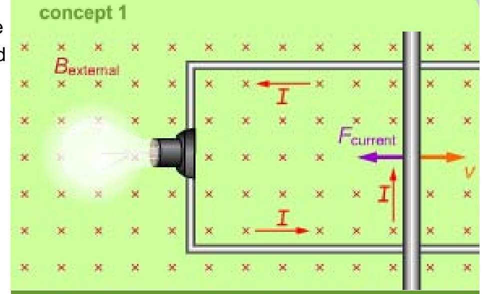
32.12 - Motion, forces and induced magnetic fields

Moving a wire through an external magnetic field induces an emf that can drive a current in a circuit. In this section, we discuss the direction of the force required to move the wire, the direction of the induced current, and the orientation of the resulting induced magnetic field.

A force is required to cause the wire to move at a constant velocity. This force counters the force that the magnetic field exerts on the moving charges that constitute the induced current. Specifically, the force F required to move the wire at a constant velocity is equal in magnitude and opposite in direction to the force F_{current} exerted on the current-carrying wire by the magnetic field. When the wire slides to the right, as shown in Concept 1, F_{current} must point to the left.

The orientation of F_{current} described above can be explained by energy considerations. Since charge carriers moving through the wire must be traveling either up or down, F_{current} must point either right or left. If F_{current} pointed in the same direction as the wire's velocity, then it would accelerate the wire. Increasing the velocity would generate even more current, more force, and so on, creating an infinite source of energy. Alas, there is no such source, so F_{current} must point to the left, opposing the force F that is pushing the wire to the right.

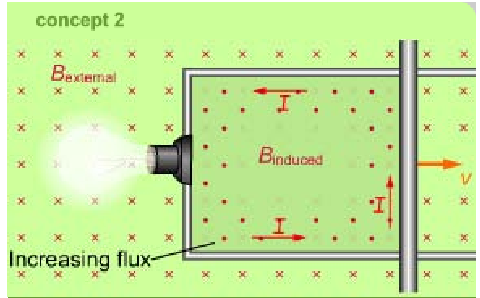
concept 1



Motional induced emf
Current in magnetic field experiences force
Force on current opposes wire's motion
Right-hand rule indicates direction of current

Since the direction of F_{current} and the magnetic field are both known, the direction of the current can be determined using the right-hand rule for moving charges in a magnetic field. The thumb needs to end up pointing in the direction of F_{current}, and the fingers must wrap from the direction of the current to the direction of the external magnetic field. In the illustration to the right, this means that the current flows up, toward the top of the screen.

This scenario can also be viewed as a case of increasing flux, which gives us an opportunity to apply Lenz's law. The area enclosed by the moving wire together with the other wires increases over time when the wire moves to the right. This means that the flux, the product of the magnetic field and the surface area enclosed by the wires, also increases.



Applying Lenz's law
Induced magnetic field due to current opposes change in flux

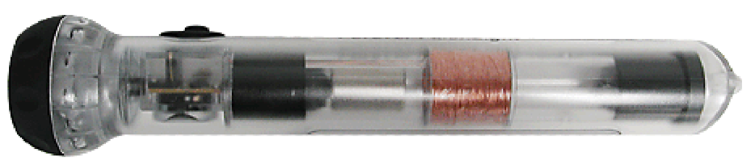
Lenz's law states that the magnetic field of the induced current will oppose the change in flux, which in this case means that inside the circuit loop it points in the direction opposite to the external field. This opposes the increase in flux due to the increase in surface area. The right-hand rule can then be used to determine the direction of the current. The induced magnetic field points toward you inside the rectangle, opposite to the external field. This means that the current flows counterclockwise, the same conclusion we reached above by using energy to analyze the directions of the forces on the moving wire segment.

The analysis using Lenz's law is summarized in Concept 2, where you see the induced magnetic field of the current pointing toward you inside the rectangle, and the current flowing counterclockwise. (For the sake of visual clarity, we dim the external field inside the loop so that you can more clearly see the field created by the induced current.)

If the amount of flux were decreasing − if the wire were sliding to the left − then the magnetic field of the induced current would point in the same direction as the external field. This orientation would oppose the **reduction** in flux arising from the decrease in the amount of surface area. In that case, the current would flow clockwise.

32.13 - Physics at work: an "induction-powered" flashlight

The principle of conservation of energy applies to an "induction-powered" circuit. Above, you see a flashlight that has no batteries but uses induction instead to transform kinetic energy into electric energy. Shaking the flashlight moves a magnet back and forth through a wire coil of many loops. The induced current charges a capacitor, which functions analogously to a battery, providing a continuing source of energy. We place "induction-powered" in quotes because the actual source of the power is the motion of the hand shaking the flashlight. Induction is used to transform this energy into a form that can be used by the light bulb.

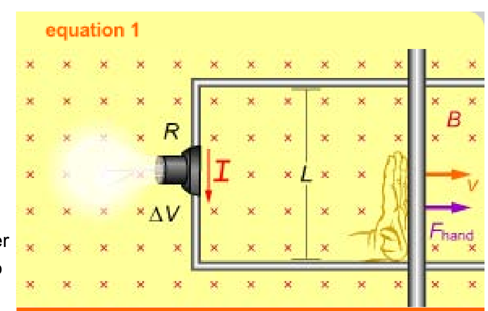


After being shaken, this "induction-powered" flashlight shines for 5 minutes. The induced current stores charge in a capacitor.

To apply energy (and power) considerations to an "induction-powered" circuit, we will use the example of a wire segment moving perpendicularly to a magnetic field, as shown in Concept 1. In this scenario, a hand pushes the wire at a constant velocity through a magnetic field of uniform strength. The circuit includes a light bulb, which we use to represent the energy-dissipating component in the circuit. Moving the wire through a magnetic field induces an emf that causes a current. This current lights the bulb, and the power consumed by the bulb equals the current times the potential difference across it.

In the steps below, we want to relate the power consumed by the light bulb to the power required to slide the wire through the magnetic field at a constant velocity. It is useful to think in terms of power (energy per unit time) since the induced emf is proportional to the change of magnetic flux per unit time.

We will find that the power supplied by the hand equals the power dissipated by the light bulb, as should be expected since that accords with the conservation of energy. The derivation also confirms statements made in a previous section about the force required to slide a wire through a magnetic field.



In an "induction-powered" circuit
Power dissipated by bulb is

$$P = \frac{L^2 v^2 B^2}{R}$$

Power expended by hand is

$$P = \frac{L^2 v^2 B^2}{R}$$

Variables

induced current in circuit	I
magnitude of induced emf in circuit	$\mathcal{E}$
potential difference across light bulb	ΔV
resistance of bulb	R
length of sliding wire	L
speed of sliding wire	v
strength of external magnetic field	B
power consumed by bulb	P
force exerted by hand pushing wire to right	F_{hand}
time interval	Δt
distance wire moves during time Δt	Δx
work done on wire to move it distance Δx	W

Energy is conserved

Strategy

1. Determine the current flowing through the light bulb in terms of the potential difference across the bulb and its resistance.
2. Determine the magnitude of the emf induced in the moving wire; it equals the amount of potential difference across the light bulb.
3. Next, determine the power consumed by the bulb. Since the power equals current times potential difference, use the equations just developed to replace the factors for current and potential difference.
4. Then, determine how much power is put into this system by the hand pushing the wire. This power equals the work it takes to slide the wire, per unit time.

Physics principles and equations

We use Ohm's law.

$$\Delta V = IR$$

The equation for the magnitude of the induced emf in a wire moving perpendicularly to a magnetic field is

$$\mathcal{E} = LvB$$

and according to Kirchhoff's loop rule this magnitude also equals the potential difference across the bulb,

$$\mathcal{E} = \Delta V$$

The power consumed by the light bulb is

$$P = I\Delta V$$

The work done by the hand pushing the wire is

$$W = F_{\text{hand}}\Delta x$$

and the power it expends is

$$P = W/\Delta t$$

The force on a current-carrying wire in a magnetic field, when the wire is perpendicular to the field, is

$$F_{\text{current}} = ILB$$

Step-by-step derivation

First, we find the potential difference across the light bulb, and the amount of current flowing through it.

Step	Reason
1. $I = \Delta V / R$	Ohm's law
2. $\Delta V = \mathcal{E} = LvB$	Kirchhoff's loop rule, motional induction
3. $I = \dfrac{LvB}{R}$	substitute equation 2 into equation 1

Now we calculate the power dissipated by the light bulb.

Step	Reason
4. $P = I\Delta V$	electric power
5. $P = \left(\frac{LvB}{R}\right)LvB$	substitute equations 2 and 3 into equation 4

Now we calculate the power in a different way, by finding out how much work the hand pushing the wire does per unit time. The result is the same equation we derived immediately above, proving that the power supplied by the hand equals the power consumed by the light bulb.

Step	Reason
6. $W = F_{\text{hand}}\Delta x$	definition of work
7. $P = \frac{W}{\Delta t} = \frac{F_{\text{hand}}\Delta x}{\Delta t}$ $P = F_{\text{hand}}v$	definitions of power, work, and speed
8. $F_{\text{hand}} = ILB$	magnetic force on wire
9. $P = (ILB)v = I(LvB)$	substitute equation 8 into equation 7 and rearrange
10. $P = \left(\frac{LvB}{R}\right)LvB$ $P = \frac{L^2v^2B^2}{R}$	substitute equation 3 into equation 9

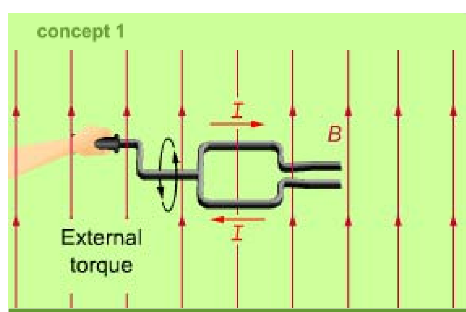
The equation in step 5, derived by considering the induced emf, is the same as the equation in step 10, derived by considering the force required to push the wire. This shows that the power expended by the hand pushing the wire through the magnetic field equals the power dissipated by the light bulb in the circuit.

32.14 - Electric generators

This bank of eight huge hydro-powered electric generators in the Grand Coulee dam produces enough electricity to supply the needs of a large city.

A loop of wire rotating in a magnetic field can be used as an electric generator to create electric current. Although the loop generally moves at a constant angular velocity, the angle it makes with the magnetic field is constantly changing. You have already seen examples of induction resulting from changing field strength and changing area. Generators provide an example of electromagnetic induction arising from a changing angle, the third factor affecting flux.

In this section, we show a hand crank turning a single generator loop, which means a human supplies the energy to keep it turning. In actual power plants, the energy might come from the burning of coal or from water falling through a turbine in a dam. The term "generator" is something of a misnomer, since this equipment actually transforms one type of energy to another rather than generating it from scratch.



Electric generator
Converts mechanical to electrical energy

There are two views of the generator shown in Concepts 2 and 3. In these diagrams, the view on the right shows the generator rotated 90° from the view on the left. In the right-hand view, you are viewing the generator from behind the crank, but the crank has been omitted to simplify the diagram. We also do not show the rest of the circuit of which the loop is a part.

As the wire loop turns, its lengthwise (horizontal) segments move through the magnetic field, which induces an emf and causes a current to flow in them. In many earlier examples, we considered a wire moving perpendicularly to a magnetic field; here, each segment moves in varying directions through the field and we must account for this. Remember that the amount of force acting on a moving charge in a magnetic field is the product of the charge, its speed, the strength of the magnetic field, and the sine of the angle between the charge's velocity and the field. The amount of induced emf in each segment changes over time because the angle between its velocity and the magnetic field changes as the generator loop rotates.

The induced emf is at its maximum when the lengthwise segments move perpendicularly to the field. In the illustration to the right, this would be when the loop is vertical. The induced emf is zero when the segments move parallel to the field, which would be when the loop is horizontal.

The illustration in Concept 3 shows the direction of a force caused by the motion of the current-carrying wire through the magnetic field. The existence of this force creates a torque that is called a *countertorque*. The person turning the crank supplies the torque to match this countertorque. Again, two views of the countertorque are shown, one facing the wire loop, and one from behind the crank. The "behind" view emphasizes that the forces causing the countertorque point in opposite directions on the two lengthwise segments of the loop. This is because the current flows in one direction on one side of the loop, and in the opposite direction on the other side.

The orientation of the emf, and the direction of the current, alternate as the crank turns. Why? As it completes a full cycle of movement, each lengthwise segment spends half its time moving in one direction relative to the field, and half the time moving in the other.

This means the orientation of the emf and the direction of the induced current change every half cycle. In addition, the magnitude of the emf changes as the wires move. This magnitude is a function of the sine of the angle between a lengthwise segment's velocity and the field. When the angle equals $90°$, the induced emf is at its maximum; when it equals $0°$, the induced emf is zero.

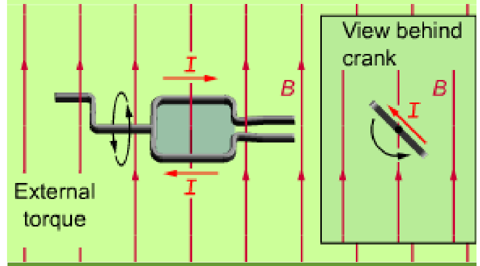
We have chosen to model this simple generator as two parallel wires moving through a magnetic field. Another way to model it would be to consider the loop as enclosing a surface area. As the loop rotates, the angle between the loop's area vector and the magnetic field changes. We could determine how the flux changes over time due to this change in angle. The description of the induced emf and current arising from such a model would be exactly the same as the one obtained from the parallel-wires model.

The generator creates an *alternating current* (AC). This current alternates in direction and varies in magnitude for the reasons mentioned above. Standard electrical outlets in homes, schools and businesses supply alternating current, current that changes cyclically in direction and magnitude, completing 60 cycles each second.

The 120-volt rating of an electrical outlet expresses a form of its average emf during a complete cycle. Its maximum emf is about 170 V. The actual average varies by time and location, and can get as low as 110 V without adversely affecting the operation of appliances. The standard AC in many European countries oscillates at 50 cycles per second, with an average emf of 220-230 V.

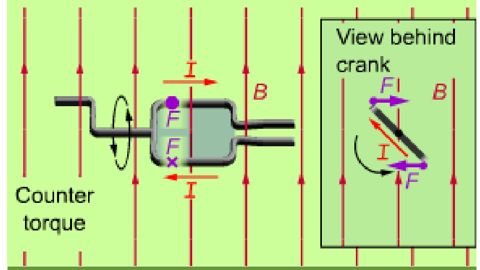
Concept 4 shows the symbol that is used in circuit diagrams to represent an AC generator.

concept 2



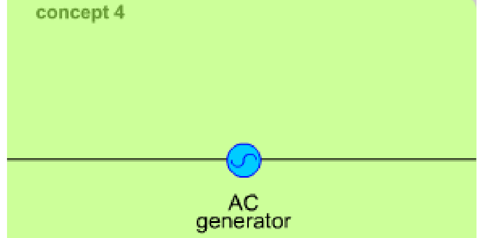
Electric generator
Wire moves in magnetic field
emf induced in loop
· Depends on angle between motion and field

concept 3



Countertorque
Magnetic force on loop creates countertorque

concept 4



Alternating current (AC) generators
Symbol for AC generator

32.15 - Interactive problem: a generator

In this simulation your task is to establish how the angle between a magnetic field and the velocity vector of a horizontal segment of a rotating loop relates to the emf induced in the loop.

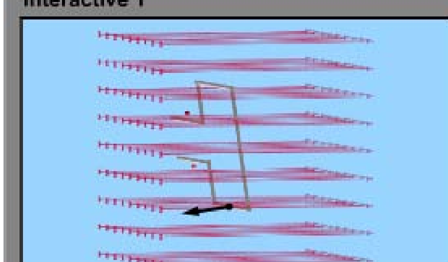
When you open the simulation by clicking on the diagram to the right, you will see a loop of wire suspended in a magnetic field. Start the loop spinning by pressing GO.

An oscilloscope displays the emf induced in the loop. For simplicity's sake, we omit showing the oscilloscope's connection to the generator circuit. We also show the current induced in the visible portion of the circuit.

A black vector represents the tangential velocity of a horizontal loop segment. You control how fast the loop rotates by setting its angular velocity. You can also change the angle at which you view the simulation. The side view of the magnetic field is particularly helpful for this exercise.

Watch the loop spinning in the field. Observe the orientation of the velocity vector to the magnetic field lines. (The angle θ displayed in the control panel is the angle

interactive 1

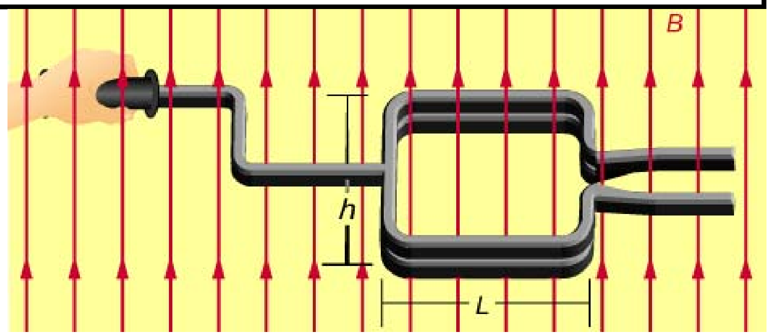


Observe amount of emf induced in the spinning generator coil between the field and the velocity vector as viewed from behind the crank.) With what orientation of the velocity vector is the emf a maximum? A minimum? Can you explain why?

Also, observe the oscilloscope trace of the emf as the generator spins faster or slower. This always appears to be a sinusoidal wave. Does its frequency change with angular velocity? Does its amplitude? To answer the second question, consider the relationship between speed and the induced emf when a wire moves in a straight line through a magnetic field.

If you need help answering these questions, review the section of the textbook on electric generators and the section on motional induction.

32.16 - Quantifying the emf induced in a generator



Hand cranked generator with a coil consisting of two loops.

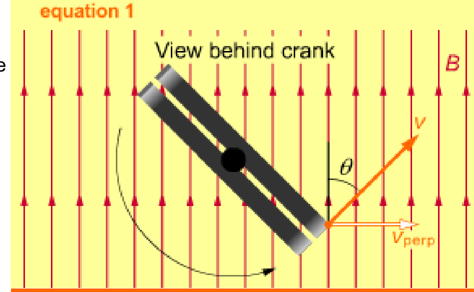
In Equation 1, you see the equation used to calculate the emf induced as the generator above is rotated in a magnetic field. The generator will produce an alternating current. The current and the induced emf vary sinusoidally over time. As you might expect, the greater the number of loops, the stronger the magnetic field, the larger the generator coil, or the faster it is rotated, then the greater the amplitude of the induced emf.

The coil is assumed to rotate through a uniform magnetic field $\mathbf{B}$ at a constant angular velocity ω. It consists of N rectangular loops of wire, each having length L and width h. Although the angular velocity is constant, the angle θ between the field and the velocity vector $\mathbf{v}$ of a lengthwise loop segment varies, as shown in the "View behind crank" illustration on the right. In this view we also show the component v_{perp} of the velocity vector that is perpendicular to the field.

The expression for the emf induced in the coil of this generator is derived in the steps below. To analyze the emf induced in the entire loop, we consider in turn the emf induced in each of its segments. "Upper" and "lower" lengthwise segments are moving in opposite directions at any moment in time. For instance, in the illustration above, the upper segment might be moving down and away from you while the lower is moving up and toward you. Given the directions of their motions relative to the field, their induced emfs add so that their combined emf in the generator circuit equals twice the emf of either segment alone.

"Left" and "right" crosswise segments (of height h) move in the same direction through the field at any instant and their emfs have a parallel orientation that causes them to cancel in the generator circuit. This means we do not have to further concern ourselves with the "crosswise" emfs in this derivation.

equation 1



Induced emf in a generator

$$\mathcal{E} = NBA\omega \sin \omega t$$

$\mathcal{E}$ = emf
N = number of loops
B = magnetic field strength
A = area of one loop
ω = angular velocity
t = time

Variables

induced emf in segment	$\mathcal{E}$
magnetic field strength	B
length of rectangular loop in coil	L
width of rectangular loop in coil	h
area of rectangular loop in coil	A
speed of lengthwise segment	v
component of speed perpendicular to field	v_{perp}
distance from axis of rotation to sides of loop	r
angle between tangential velocity and field	θ
angular velocity of rotation	ω
time	t
number of loops in generator coil	N

Strategy

1. Write the equation for the induced emf in a wire that moves through a magnetic field. Use trigonometry to describe the component of the motion that is perpendicular to the field.
2. Express the perpendicular component in terms of the angular velocity ω of the generator coil.
3. Express the emf induced in one lengthwise segment of the generator coil in terms of ω.
4. Calculate the emf induced in the entire coil.

Physics principles and equations

The equation for the magnitude of the emf induced in a wire moving perpendicularly with speed v_{perp} through a uniform magnetic field is

$$\mathcal{E} = Lv_{perp}B$$

Each horizontal segment of a generator loop moves in a circular path through the magnetic field with velocity $\mathbf{v}$, resulting in an emf that alternates in direction as the segment completes each half cycle. The diagram in Equation 1 shows that the component of $\mathbf{v}$ perpendicular to the field is $v_{perp} = v \sin \theta$. The emf for a wire segment equals

$$\mathcal{E} = L(v \sin \theta)B = BLv \sin \theta$$

Equations for tangential speed and angular velocity are:

$$v = r\omega, \quad \omega = \frac{\theta}{t}$$

Step-by-step derivation

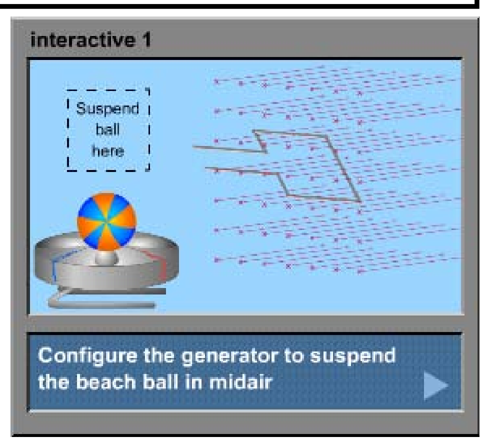
We start with the equation for the induced emf in a wire that moves at an angle in a magnetic field. Then we perform several substitutions to account for the geometry and angular velocity of the generator coil.

Step	Reason
1. $\mathcal{E} = BLv \sin \theta$	emf of wire moving at angle through magnetic field
2. $v = r\omega = ½h\omega$	equation for tangential speed
3. $\theta = \omega t$	definition of angular velocity
4. $\mathcal{E} = BL(½h\omega) \sin \omega t$	substitute equations 2 and 3 into equation 1

Knowing the emf generated in each lengthwise segment, we add them all up to find the total emf $\mathcal{E}_{total}$ induced in the generator coil.

Step	Reason
5. $\mathcal{E}_{total} = 2N\mathcal{E}$	total emf from N two-sided loops
6. $\mathcal{E}_{total} = NB(Lh)\omega \sin \omega t$	substitute equation 4 into equation 5
7. $\mathcal{E}_{total} = NBA\omega \sin \omega t$	area of rectangle

32.17 - Interactive problem: configuring a generator



In this interactive simulation, your task is to configure a generator to produce an emf that has just enough amplitude to power an electric fan.

A beach ball sits atop the fan. You want to power the fan to elevate the beach ball to the height indicated on the screen. To do this, you need to generate an emf with a maximum value of 168 volts across the fan's electrical circuit. (We say "maximum value" since the emf will vary sinusoidally between positive and negative 168 volts.) In the simulation, you can set the following variables:

1. The whole number of loops of wire in the generator coil, from one to four.
2. The magnitude of the magnetic field through which the coil spins, from 0.00 to 6.00 T in increments of 1.00 T.
3. The loop area, from 0.20 to 2.00 m^2 in increments of 0.10 m^2.
4. The angular velocity ω of the coil, from 0.00 to 6.00 rad/s in increments of 0.10 rad/s.

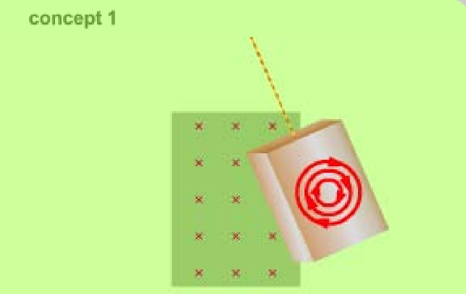
Configure the generator to induce a maximum emf of 168 volts. You will have to adjust all the values from their initial settings to induce the correct maximum emf. When you think you have it right, press GO to see if the beach ball rises to the target height.

Too much emf and the ball may fly off the screen. Too little and it will never make it into the air. It may prove helpful to use trial and error to determine approximately correct values before using the equation to solve the problem.

An oscilloscope displays the emf created by the generator. It provides another way for you to judge the success of your efforts.

For assistance with this problem, review the section of the textbook on quantifying the emf induced by a generator.

Eddy currents: Electric currents created when a solid conductor moves through a magnetic field.

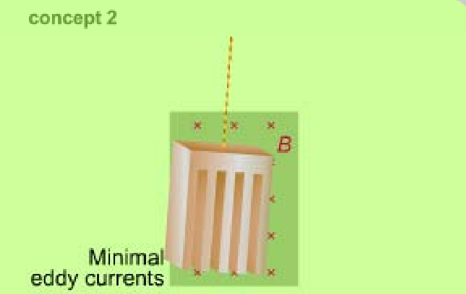


Eddy currents
Caused by changing flux in solid conductors
Eddy currents damp motion

We have used a relatively simple configuration to analyze the current created by moving a wire in a magnetic field. The moving wire was a straight segment, and an overall rectangular circuit contained three other straight segments.

Here, we use eddy currents to illustrate a significant point: Even when a circuit is not well defined, currents can still be caused to flow in a conductor at rest in a time-varying magnetic field or passing through a nonuniform magnetic field. For instance, the rectangular block to the right is made of a conducting material. When it moves into or out of the magnetic field shown, currents flow in it, but not in straight lines. These types of current flow are called *eddy currents*, and they flow in a complex pattern that resembles a whirlpool. The diagram in Concept 1 visually approximates the eddy currents by depicting them as circular.

The existence of eddy currents can be vividly demonstrated with two differing pieces of conducting material. One piece is a solid block attached to a string, swinging in and out of a magnetic field. This motion generates eddy currents in the block. The magnetic fields of the eddy currents oppose the change in flux that causes them. In essence, the block becomes an electromagnet whose motion is opposed by the external magnetic field. The opposing force quickly damps (reduces) the motion.



Slotted block
Slots minimize eddy currents
Oscillation minimally dampened

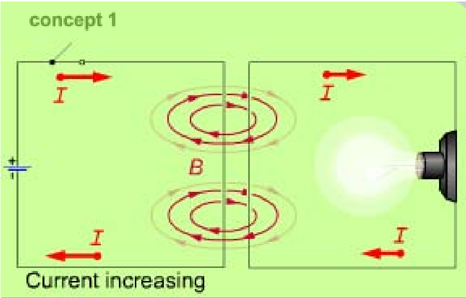
A second block, with slots cut through it, swings through the same field. Because the slots prevent the formation of significant large-scale eddy currents, a much weaker opposing magnetic field is created. There is less force to damp the motion, and the block swings relatively freely.

Like any phenomenon that removes mechanical energy from a system, causing it to dissipate as heat, eddy currents can be either an undesirable source of inefficiency, or in the right circumstances, a valuable tool. One useful application of eddy currents is to the braking of trains.

If the upper half of a spinning train wheel is subjected to a strong magnetic field, each portion of the wheel will rotate into and out of this field. The changing magnetic flux in that portion of the wheel will induce eddy currents that apply a countertorque to the whole wheel, acting to slow it down. Since the magnitude of the countertorque is proportional to the angular velocity of the wheel, the braking effect diminishes as the train slows, allowing it to come to a smooth stop. Eddy-current brakes are promoted as being quieter and less prone to slippage than train brakes that rely on mechanical friction.

32.19 - Mutual induction

Mutual induction: The induction of an emf in one circuit by a changing current in another circuit.



Mutual induction
Changing current in one circuit creates current in other
· Changing current causes changing magnetic field
· Changing magnetic field induces current

Currents create magnetic fields. A changing current in a circuit creates a changing magnetic field. This changing magnetic field will induce an emf and a current in a nearby second circuit, in a process called mutual induction.

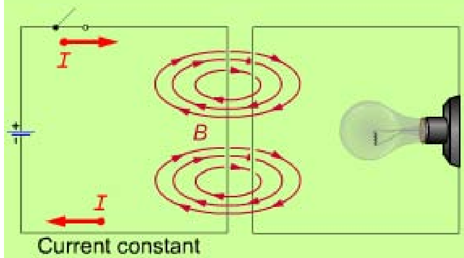
At the right, you see an illustration of the steps involved in mutual induction:

1. A switch is closed in the left-hand circuit. This circuit contains a battery.
2. A current in the left-hand circuit starts to flow after the switch is closed, increasing toward its steady state value.
3. The increasing current on the left generates a changing magnetic field.
4. The changing magnetic field causes a changing magnetic flux through the circuit on the right, which induces an emf in this circuit.
5. The induced emf in the right-hand circuit causes a current there, lighting the light bulb for a brief time.
6. Once the current on the left reaches its steady state, the light bulb on the right goes out, because there is no longer a changing magnetic field to induce a current there.

An important point to remember is that currents are induced by *a* **changing** magnetic field. The closing of the battery circuit starts a current that in turn creates a magnetic field where before there was none. It is this **change** in magnetic field that induces the emf and the current in the second circuit. When the current caused by the battery reaches a steady state, which occurs quite quickly, then its magnetic field will also be constant and it will cease to induce an emf or a current in the circuit on the right.

If the current on the left-hand side were an alternating instead of a direct current, it would induce an undiminishing alternating current on the right-hand side. Why? The magnetic field due to the changing left-side current would constantly change in strength and orientation. The constantly changing magnetic field would continuously induce an emf and current in the circuit on the right, and the light bulb would stay lit.

concept 2



Mutual induction
When current is constant, no induction occurs

32.20 - Transformers

Transformer: A device used to increase or decrease an alternating potential difference.

Power transmission lines conduct AC current at high voltages. Transformers convert this current from one voltage to another.

The way in which electricity is transmitted and distributed requires conversions of potential differences across an enormous range of values. Long distance power transmission lines typically operate at about 350,000 V, while city power lines near a home function at 15,000 V. Before the electricity enters a house, the potential difference between a current-carrying and a "neutral" wire is further reduced to 120 V. All of these currents are alternating, and transformers are used to convert potential differences in alternating currents. (One reason why alternating current is used in power systems is the ease with which transformers can change its potential difference.)

The illustration in Concept 1 shows a transformer: two coils of wire wrapped around an iron core. The wires are insulated so that no current flows directly between them. The iron core allows a magnetic field to be efficiently transmitted from one coil to the other: Less than 5% of the energy transformed in a typical transformer is dissipated as heat. In this section, we focus on the characteristics of an ideal transformer, one in which no energy is lost.

The wire wrapped around the iron core on the left-hand side is called the *primary winding*. In our scenario, it is directly attached to an AC generator. The current passing through this coil generates a magnetic field. Since the current continually changes, so does the magnetic field it creates.

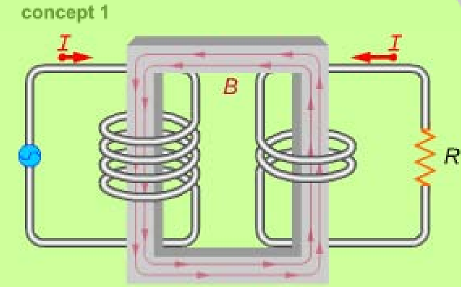
The field passes through the second coil on the right. This coil is called the *secondary winding*. The change in magnetic flux through this coil induces an emf that causes a current in the secondary circuit. The secondary circuit powers a resistive component, such as a light bulb, that is called the *load*. The transformer changes the emf from that created by the AC generator to that required by the load.

There are a different number of loops, or turns, in the primary and secondary windings. This is crucial to the operation of the transformer. The ratio of the number of primary loops to the number of secondary loops is called the *turns ratio*.

The transformer "transforms" the potential difference created by the generator: The potential difference across the secondary winding differs from that across the primary winding. In the transformer to the right, the potential difference decreases. This is a *step-down* transformer. *Step-up* transformers increase potential difference. The transformer symbol that is used in circuit diagrams is shown in Concept 2.

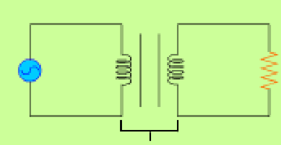
The change in the potential differences depends on the turns ratio. The proportion in Equation 1 states that the ratio of the potential differences equals the turns ratio. To put it another way: If the number of loops on a given side is reduced, the potential difference on that side decreases.

concept 1



Transformers
Increase or decrease potential difference
AC creates changing magnetic field
Induces potential difference on right

concept 2



Transformer
Symbol used in circuit diagrams

We derive this "turns equation" below using Faraday's law. The loops in each coil are assumed to span an equal surface area, and the same amount of magnetic field passes through each one. This means the ratio of the changes in flux in the primary and secondary windings is proportional solely to the number of loops in each coil.

Variables

potential difference across primary coil	ΔV_1
potential difference across secondary coil	ΔV_2
number of loops in primary coil	N_1
number of loops in secondary coil	N_2
time interval	Δt
change in magnetic flux through one loop during Δt	$\Delta \Phi_B$

Strategy

1. State Faraday's law twice, for the emf induced across each coil by the changing magnetic flux. The emf equals the potential difference across the components.
2. Divide the two equations.

Physics principles and equations

Faraday's law stated in terms of potential difference, and for a coil of N loops, is

$$\Delta V = -N\frac{\Delta \Phi_B}{\Delta t}$$

With transformers, it is traditional to refer to potential differences rather than emfs in the primary and secondary windings.

Step-by-step derivation

We state Faraday's law for each coil, then divide the two equations and simplify the result.

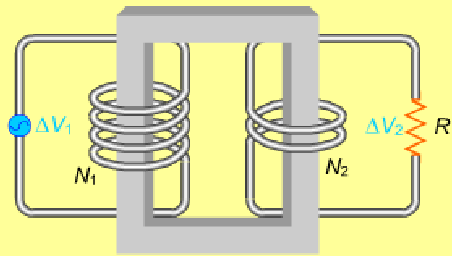
Step	Reason
1. $\Delta V_1 = -N_1\frac{\Delta \Phi_B}{\Delta t}$	Faraday's law
2. $\Delta V_2 = -N_2\frac{\Delta \Phi_B}{\Delta t}$	Faraday's law
3. $\frac{\Delta V_1}{\Delta V_2} = \frac{-N_1\frac{\Delta \Phi_B}{\Delta t}}{-N_2\frac{\Delta \Phi_B}{\Delta t}}$	divide
4. $\frac{\Delta V_1}{\Delta V_2} = \frac{N_1}{N_2}$	simplify

Transformers can increase a potential difference, but by now you have learned that in physics, you do not get something without also giving something up. The principle of conservation of energy must apply to transformers. How does this precept apply? It requires that the power, or energy produced or consumed per unit time, is the same on both sides of an ideal transformer. Power equals the current times the potential difference. If you boost the potential difference, you decrease the current.

The ratio of the currents is the inverse of the turns ratio. This is shown in Equation 2 .

The example problem asks you to analyze a type of transformer commonly sold in travel stores. This transformer steps down European voltage (240 V AC) to USA standard voltage (120 V AC), and you may need one if you intend to use any of your personal appliances on a trip abroad.

equation 1



Potential differences ratio equals turns ratio

$$\frac{\Delta V_1}{\Delta V_2} = \frac{N_1}{N_2}$$

ΔV_1 = primary potential difference
ΔV_2 = secondary potential difference
N_1 = primary number of loops
N_2 = secondary number of loops

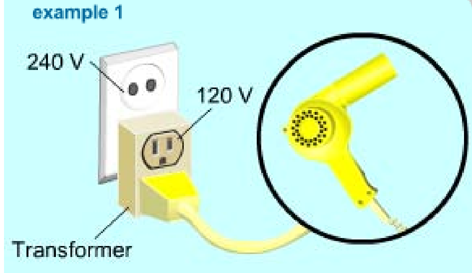
equation 2

Currents ratio is inverse of turns ratio

$$\frac{I_2}{I_1} = \frac{N_1}{N_2}$$

I_1 = primary current
I_2 = secondary current

example 1

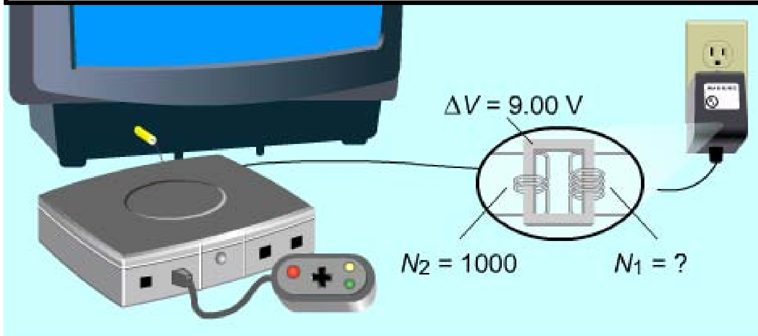


What should the turns ratio of the transformer be?

$$\frac{\Delta V_1}{\Delta V_2} = \frac{N_1}{N_2}$$

$$\frac{240\text{ V}}{120\text{ V}} = \frac{N_1}{N_2} = 2:1$$

32.21 - Sample problem: primary winding in a transformer



A video game console operates at a potential difference of 9.00 V. The potential difference supplied by the power receptacle is 120 V AC. A transformer is used to convert one voltage to the other. If the secondary coil of the transformer (the 9.00 V side) has 1000 loops, how many loops does the primary coil have?

Low-voltage devices like the game console above typically run on direct current, so the power converter would not only transform the potential difference, but also convert AC to DC. Here, we only consider the AC transformer part of the power converter.

Variables

primary potential difference	$V_1 = 120$ V
secondary potential difference	$V_2 = 9.00$ V
primary number of loops	N_1
secondary number of loops	$N_2 = 1000$

What is the strategy?

1. State the turns ratio equation.
2. Solve for the number of primary loops and evaluate.

Physics principles and equations

The turns equation for transformers is

$$\frac{\Delta V_1}{\Delta V_2} = \frac{N_1}{N_2}$$

Step-by-step solution

We solve the turns equation for N_1, then evaluate.

Step	Reason
1. $\frac{\Delta V_1}{\Delta V_2} = \frac{N_1}{N_2}$	turns equation
2. $N_1 = \frac{\Delta V_1}{\Delta V_2} N_2$	solve for number of primary loops
3. $N_1 = \frac{120\text{V}}{9.00\text{V}}(1000)$ $N_1 = 1.33 \times 10^4$	evaluate

32.22 - Interactive problem: configuring a transformer


The large cylindrical transformers on the left of this photo convert electricity from 350,000 V to 15,000 V at an electrical substation.

Here, we simulate a transformer.

An alternating potential difference across the primary winding on the top creates a magnetic field that passes through the iron core to the secondary winding on the bottom. This magnetic field induces an emf in the bottom coil. The value of the potential difference across the resistor in the bottom, secondary circuit (equal to the induced emf) is displayed on an oscilloscope. You can turn the dial on the oscilloscope to change the scale of the display. The height of a grid box equals the setting on the dial.

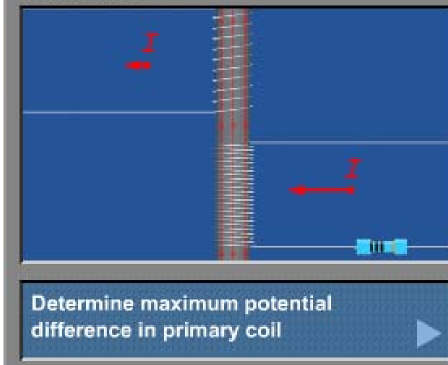
Your task is to determine the maximum positive potential difference in the top, primary circuit using information you glean from the simulation. We use the word “maximum” because the potential difference of the alternating current varies sinusoidally with time.

There are ten loops in the top coil. You can vary the number of loops in the bottom coil from 20 to 50, in increments of 5. Pick a number of loops and then use the oscilloscope to determine the maximum potential difference across the resistor in the secondary circuit. You can use that value and the ratio of loops to determine the maximum potential difference in the primary circuit.

Enter this value in the box provided in the control panel and press CHECK to see if you are right. If not, redo your calculations, enter a new answer, and press CHECK again.

If you need help, review the section on transformers.

interactive 1



32.23 - Self-induction

Self-induction: The emf induced in a circuit by a changing current in that circuit.

Transformers exploit mutual induction. The same effect can occur within a single component in one circuit, as well. This phenomenon is called self-induction. It is usually discussed in the context of an *inductor*, which is essentially a tightly wrapped coil of wire. (The diagram on the right shows a loosely wrapped coil so that you can better see the loops that make it up.)

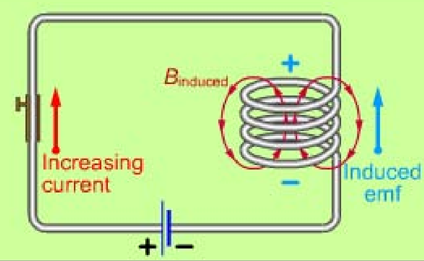
Self-induction occurs for the same reason that mutual induction does: A changing current causes a changing magnetic field, which in turn induces an emf. The induced emf obeys Lenz's law, which means it acts to counter the **change** that created it. The induced emf opposes any change in current, whether it is an increase or a decrease.

We illustrate this on the right, using the example of a circuit in which the switch has just been closed. After the switch closes, the current begins to increase. This means the magnetic field it causes in the coil is increasing, and Lenz's law dictates that the **induced** magnetic field, which is shown in Concept 1, must be oriented to oppose this increase. The emf induced in the inductor must oppose that of the battery. We show the orientation of the induced emf in the inductor using + and − signs, and you see that they are oriented oppositely in the circuit to those of the battery.

Our illustration in Concept 1 depicts the first few instants after the switch closes in the circuit. During this brief time period we model the current as increasing at a linear rate so that the induced magnetic field and the induced emf are constant. Over a longer time period the current would level off at a steady value determined by the resistance in the circuit and the emf of the battery, and the induced magnetic field and emf would decrease to zero. The induced emf in this circuit opposes the source emf, but it is never strong enough to nullify it.

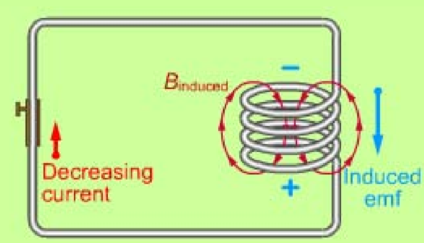
In Concept 2, we remove the battery from the circuit and the current decreases rapidly, but not instantly, to zero. When the current is decreasing, self-induction also occurs. In this case both the induced magnetic field and the induced emf are reversed, since the emf opposes the change in current that causes it. The presence of the inductor slows the decrease in the current.

concept 1



Self-induction
Changing current causes a changing magnetic field
· inducing an emf in the same wire
· which acts counter to the change in current

concept 2



Self-induced emf
Acts to oppose an increasing current
Acts to strengthen a decreasing current

Self-induction is one reason why you are advised not to pull out the plug of an electrical appliance before it is turned off. The abrupt change in current caused by yanking out the cord can cause a significant induced emf that opposes the decrease in current. This induced emf can cause sparks, as well as an unexpected and perhaps damaging current in the appliance.

32.24 - Inductors and inductance

Inductor: An electrical circuit component that has a large inductance. A change in current will induce a relatively large opposing emf in an inductor.

Inductance: Relates the induced emf of an inductor to the rate of change of current.

At the right, you see a circuit containing an inductor, which in this case is a wire coil in the shape of a solenoid. When the current in the inductor changes, it creates a magnetic field that also changes. That changing magnetic field in turn induces an emf across the inductor. The induced emf resists the change in current, causing the current to increase or decrease more slowly.

Inductance relates the amount of induced emf to the rate of change of current in a given inductor. The letter L represents the inductance of an inductor. The coil graphic appearing in each illustration to the right is the symbol for an inductor in a circuit diagram.

In Equation 1, we write the definition of inductance: It equals the induced emf divided by the rate of change of the current. In Equation 2, we write an equation for the induction in terms of magnetic flux. It, too, can serve as the definition of inductance. The henry (H) is the unit for inductance. One henry equals one tesla·square meter per ampere, or one weber per ampere. (The plural of henry is henrys.)

The first equation is highly useful in the practical analysis of circuits, where emf and a changing current are more easily measured than magnetic flux. It is often rearranged to state that the induced emf equals the negative of the product of the inductance times the rate of change of current, as you also see in Equation 1. This form of the equation shows how the induced emf increases with inductance and the rate of change of current. The negative sign reflects the fact that the induced emf opposes the change in current that caused it. The greater the inductance, the greater the induced emf opposing a given rate of change of the current.

The second equation, $L = N\Phi_B/I$, illustrates the relationship between inductance and an inductor's physical configuration. In this equation, $N\Phi_B$ is the total magnetic flux through the inductor. For a given magnetic field and current, a larger magnetic flux through the inductor is associated either with more loops (greater N) or with larger loops (greater Φ_B for each loop). In this way, inductance is a measure of how much magnetic flux passes through the inductor per ampere of current. The magnetic field and flux referred to here and in the diagrams to the right are those due to the current, not the induced magnetic field. Below, we derive the second equation from the first one, using Faraday's law.

Variables

inductance of inductor	L
induced emf	$\mathcal{E}$
current	I
time interval	Δt
number of loops in inductor	N
magnetic flux through one loop	Φ_B

Strategy

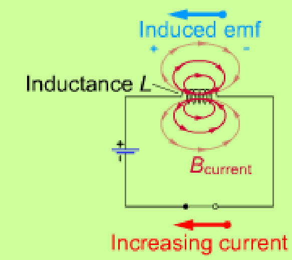
1. State the definition of inductance and substitute the expression for the induced emf from Faraday's law.
2. Simplify.

Physics principles and equations

Faraday's law stated with the number N of loops in a coil:

$$\mathcal{E} = -N\frac{\Delta\Phi_B}{\Delta t}$$

concept 1

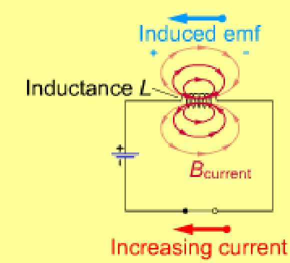


Inductors and inductance

Inductor's emf opposes change in current

Inductance relates emf to rate of change of current

equation 1

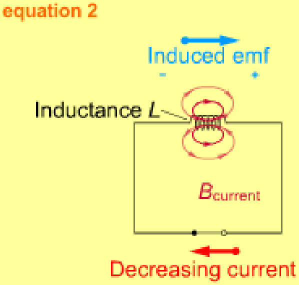


Inductance and induced emf

$$L = -\frac{\mathcal{E}}{\Delta I/\Delta t} \quad \text{and} \quad \mathcal{E} = -L\frac{\Delta I}{\Delta t}$$

L = inductance
$\mathcal{E}$ = induced emf
ΔI = change in current
Δt = time interval

equation 2



Inductance and flux

$$L = N\frac{\Phi_B}{I}$$

L = inductance
N = number of loops
Φ_B = magnetic flux through one loop
I = current

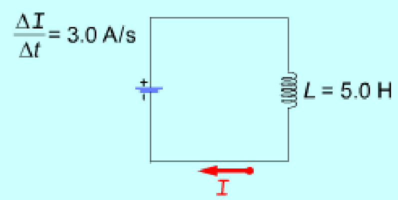
Units: henrys ($H = T{\cdot}m^2/A$)

Step-by-step derivation

We state the definition of inductance and substitute Faraday's law.

Step	Reason
1. $L = \frac{-\mathcal{E}}{\Delta I / \Delta t}$	definition of inductance
2. $\mathcal{E} = -N\frac{\Delta \Phi_B}{\Delta t}$	Faraday's law
3. $L = \frac{N\Delta \Phi_B / \Delta t}{\Delta I / \Delta t}$	substitute equation 2 into equation 1
4. $L = \frac{N\Delta \Phi_B}{\Delta I}$	simplify
5. $L = \frac{N\Phi_B}{I}$	$\Phi_B = 0$ when $I = 0$ and $t = 0$

example 1



What are the magnitude and orientation of the emf induced in the inductor?

$$\mathcal{E} = -L\frac{\Delta I}{\Delta t}$$

$\mathcal{E} = -(5.0\text{ H})(3.0\text{ A/s})$

$\mathcal{E} = -15\text{ V}$

Opposes emf of battery

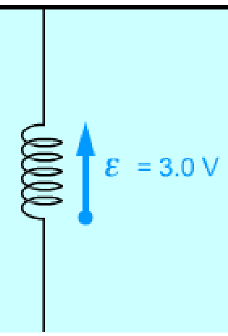
As you may have noted, the concept of inductance has some similarities to a concept you studied earlier: capacitance. Both depend on the physical configuration of their respective components. Capacitance relates potential difference and charge; inductance relates induced emf (equal to the potential difference across the inductor) and the rate of change of current.

Inductors are used in some surprising applications. Traffic engineers use large inductors at intersections as sensors to detect when cars are waiting for the traffic signal to change from red to green. They do this by burying a coil of wire from four to eight feet in diameter beneath the asphalt road surface. A current runs through this inductor.

When a car rolls to a stop over the buried coil, suddenly there is a 1200 kg metal object positioned in the coil's magnetic field. This changes the inductance of the inductor. The inductor is part of a timing circuit, and the changed inductance alters the nature of this circuit, which is interpreted as a signal in the traffic light system.

The car changes the inductance in a perhaps surprising way: It reduces it. Eddy currents flow through the car, opposing the magnetic flux through the coil and diminishing its inductance.

32.25 - Sample problem: induction



The emf induced in a 0.15 H inductor is kept at a constant 3.0 V by changing the current flowing through it.

How much time elapses as the current changes from +1.0 to −1.0 A?

Variables

inductance of inductor	$L = 0.15\text{ H}$
emf induced in inductor	$\mathcal{E} = 3.0\text{ V}$
initial current	$I_i = +1.0\text{ A}$
final current	$I_f = -1.0\text{ A}$
change in current	ΔI
elapsed time	Δt

What is the strategy?

1. Use the inductance equation stated below.
2. Solve for the time difference.
3. Evaluate the resulting equation.

Physics principles and equations

The relationship between inductance and induced emf is

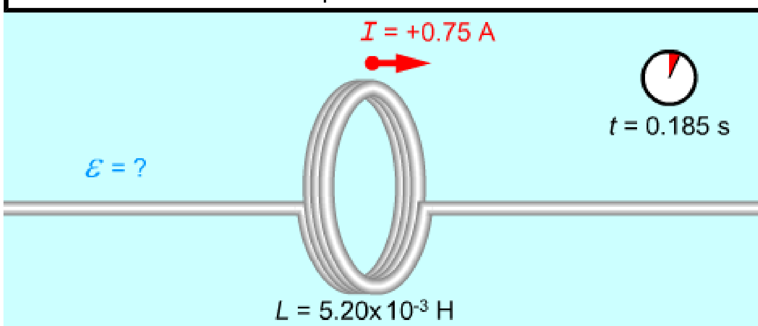
$$L = -\frac{\mathcal{E}}{\Delta I/\Delta t}$$

Step-by-step solution

We start with the definition of induction, solve for Δt, and evaluate the resulting equation.

Step	Reason
1. $L = -\frac{\mathcal{E}}{\Delta I/\Delta t}$	definition of inductance
2. $\Delta t = -L\frac{\Delta I}{\mathcal{E}}$	solve for the elapsed time
3. $\Delta I = I_f - I_i$	change in current
4. $\Delta t = -L\frac{I_f - I_i}{\mathcal{E}}$	substitute equation 3 into equation 2
5. $\Delta t = -(0.15\ \mathrm{H})\frac{-1.0\ \mathrm{A} - 1.0\ \mathrm{A}}{3.0\ \mathrm{V}}$ $\Delta t = 0.10\ \mathrm{s}$	evaluate

32.26 - Interactive checkpoint: emf in an inductor



A 5.20×10^{-3} H inductor is placed in a circuit where the current changes linearly from −1.50 A to +1.90 A in 0.250 s. What is the induced emf in the inductor during this time?

Answer:

$\mathcal{E}$ = ______ V

32.27 - Calculating the inductance of a solenoid

One common type of inductor is a solenoid, a helical coil made up of closely wrapped loops of wire. You see a photograph of solenoidal inductors above, and idealized representations of cylindrical solenoids and their magnetic fields to the right.

Although the details are rather different, the steps involved in calculating the inductance of a solenoid resemble the steps involved in calculating the capacitance of a capacitor. We compute the inductance or the capacitance based on physical characteristics of the device: area, length or separation, and so on.

These solenoidal inductors are used in electric circuits.

The inductance is proportional to the length s of the solenoid. It is also proportional to the area A of each loop, and the square of the *number density* of the loops. The number density n equals the number of loops per unit length. The inductance can be increased by adding more loops to the solenoid, by increasing the density of the loops, or by increasing the radius of the loops, which increases the surface area spanned by each one.

An actual solenoid is likely to contain a material in its core designed to enhance the magnetic field there, but for simplicity's sake it is helpful to assume that the solenoid contains only air. Air does not substantially affect the magnetic field.

In the steps that follow, we derive the equation you see to the right for the inductance of a solenoidal inductor.

Variables

We use the variable s to stand for the length of the inductor since the letter L is used for inductance.

magnetic flux through one loop	Φ_B
magnetic field strength inside solenoid	B
area spanned by one loop	A
permeability constant	$\mu_0 = 4\pi\times10^{-7}$ T·m/A
inductance of solenoid	L
current through solenoid	I
total number of loops in solenoid	N
length of solenoid	s
number density of loops	n

Strategy

1. First, calculate the magnetic flux passing through one loop of the solenoid.
2. Then use the equation for the inductance in terms of the number of loops, the flux through one loop, and the current.

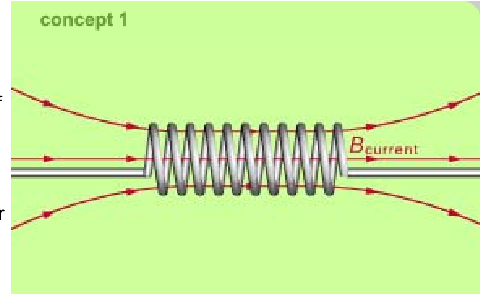
Physics principles and equations

The magnetic flux through a surface area A perpendicular to a magnetic field is

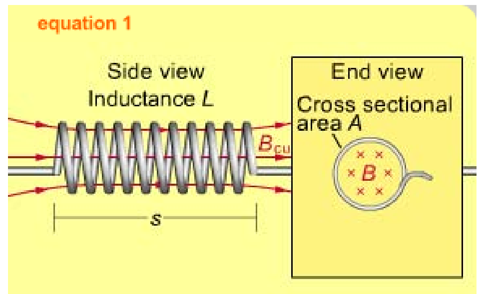
$$\Phi_B = BA$$

The strength of the magnetic field produced by a steady state current inside the solenoid, as derived in an earlier chapter, is

$$B = \mu_0 In$$



Inductors are often solenoids



Inductance of a solenoid

$$\boxed{L = n^2 s\mu_0 A}$$

L = inductance
n = number density of loops
s = length
μ_0 = permeability constant
A = loop area

Finally, we will use the equation that expresses the inductance in terms of the magnetic flux and the number of loops.

$$L = N\frac{\Phi_B}{I}$$

Step-by-step derivation

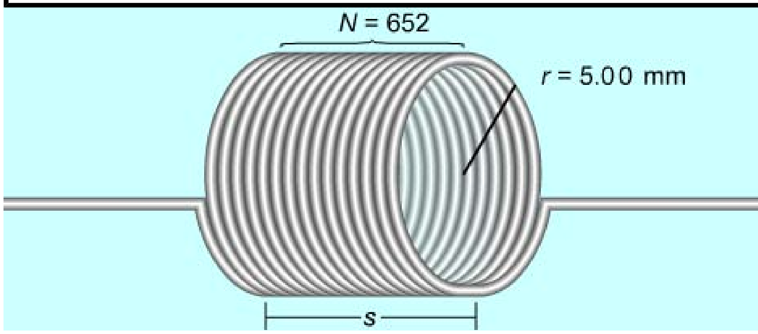
First we calculate the magnetic flux through each loop of the solenoid.

Step	Reason
1. $\Phi_B = BA$	definition of flux
2. $B = \mu_0 In$	magnetic field of solenoid
3. $\Phi_B = (\mu_0 In)(A)$	substitute equation 2 into equation 1

Now we find an expression for N, and then we substitute expressions for N and Φ_B into the inductance equation stated above. After simplifying, we get the desired formula.

Step	Reason
4. $N = ns$	total loops in terms of density
5. $L = N\dfrac{\Phi_B}{I}$	inductance equation
6. $L = (ns)\dfrac{\mu_0 InA}{I}$	substitute equations 4 and 3 into equation 5
7. $L = n^2 s\mu_0 A$	simplify

32.28 - Sample problem: configuring a solenoid



A solenoid of radius 5.00 mm and having 652 turns is used to make a 2.00×10^{-3} H inductor.

How long should the coil be?

Variables

number density of loops in coil	n
number of turns in coil	$N = 652$
length of solenoid	s
inductance of solenoid	$L = 2.00 \times 10^{-3}$ H
permeability constant	$\mu_0 = 4\pi \times 10^{-7}$ T·m/A
area of a solenoid loop	A
radius of solenoid	$r = 5.00 \times 10^{-3}$ m

What is the strategy?

1. Find the number density of the solenoid loops in terms of the total number of turns and the length of the solenoid. Also, find the loop area of the solenoid in terms of its radius.
2. Substitute the expressions for number density and area into the equation that gives the inductance of a solenoid.
3. Solve the resulting equation for the length s of the solenoid and evaluate.

Physics principles and equations

Number density is defined as

$$n = \frac{N}{s}$$

We also use the equation for the inductance of a solenoid

$$L = n^2 s \mu_0 A$$

Step-by-step Solution

In the first steps we express the number density of the solenoid loops in terms of N and s and substitute this expression into the equation for the inductance of a solenoid. We will later solve for s.

Step	Reason
1. $n = \frac{N}{s}$	number density
2. $L = n^2 s \mu_0 A$	inductance of a solenoid
3. $L = \left(\frac{N}{s}\right)^2 s \mu_0 A$	substitute equation 1 into equation 2
4. $L = \frac{N^2}{s} \mu_0 A$	simplify

Now, we write the area in terms of the radius, solve for the length s, and evaluate.

Step	Reason
5. $A = \pi r^2$	area of a circle
6. $L = \frac{N^2}{s}\mu_0(\pi r^2)$	substitute equation 5 into equation 4
7. $s = \pi\mu_0 \frac{N^2 r^2}{L}$	solve for s
8. $s = \pi(4\pi \times 10^{-7} \frac{\mathrm{T \cdot m}}{\mathrm{A}}) \frac{(652)^2(5.00 \times 10^{-3}\,\mathrm{m})^2}{(2.00 \times 10^{-3}\,\mathrm{H})}$ $s = 0.0210\ \mathrm{m} = 2.10\ \mathrm{cm}$	evaluate

32.29 - Energy stored in an inductor

The circuit on the right consists of a battery, a resistor, and an inductor. The switch has just been closed. The inductor will gain magnetic potential energy, PE_B, as the current flows. This section explains why it does so.

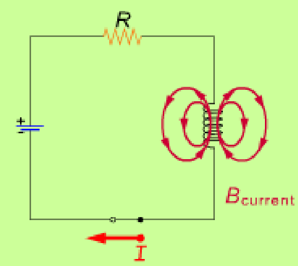
When the switch is closed, the emf of the battery causes the current to start to flow. As the current increases over time, an opposing emf is induced in the inductor in the circuit.

While the current increases, because the induced emf opposes the increasing current, work has to be done to move electrons across the inductor. This work equals the increase in the energy of the inductor, in particular, the potential energy stored in the magnetic field of the inductor.

Conversely, when the current decreases, the induced emf of the inductor will oppose that decrease, causing the current to decrease less rapidly than it would if the inductor were not present. Some of the potential energy stored in the inductor is transformed into electric energy.

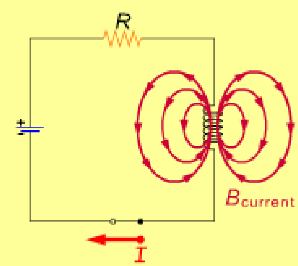
The equation in Equation 1 shows how to calculate the energy stored in an inductor. It is a function of the current and the inductance of the inductor.

concept 1



Inductors and energy
Inductor stores energy in magnetic field

equation 1

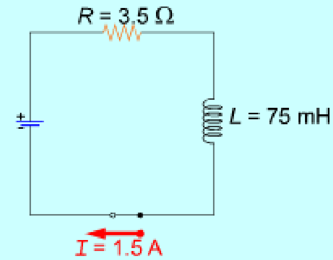


Inductors and energy

$$PE_B = \tfrac{1}{2}LI^2$$

PE_B = magnetic potential energy
L = inductance
I = current

example 1



What is the energy stored in this inductor?

$PE_B = ½LI^2$

$PE_B = ½(75×10^{-3}\text{ H})(1.5\text{ A})^2$

$PE_B = 0.084\text{ J} = 84\text{ mJ}$

32.30 - Magnetic field energy density

Magnetic field energy density (u_B): The potential energy that is stored in a magnetic field, per unit volume.

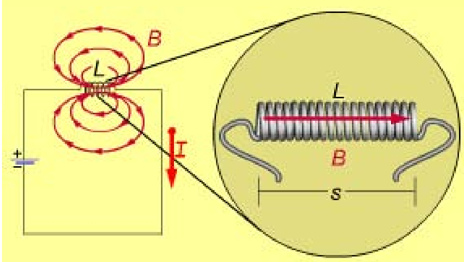
Magnetic field energy density, u_B, is the amount of magnetic energy in a magnetic field per unit volume. This definition is expressed mathematically by the first equation to the right.

If a field is uniform, then its energy density is uniform as well. If the field varies, then the energy density will vary, but the field can be considered uniform over any small enough region of space. The energy density at a point can then be calculated with the first equation by considering a very small region around that point.

The second equation states the energy density at a point in space based on the magnetic field strength there. To derive it, we use an ideal solenoidal inductor as the source of the magnetic field. The number density of the loops of the solenoid is n, its length is s, and its cross sectional area is A. Since the magnetic field produced inside the solenoid is uniform, the energy density will not vary from point to point. Although we derive the equation for magnetic field energy density using a solenoidal inductor, the equation applies to any magnetic field.

Our approach in this derivation is similar to the one we used to derive the energy density $u_E = \varepsilon_0 E^2/2$ of the electric potential energy stored in an electric field. Our use of a uniform magnetic field in an ideal solenoid in this derivation is similar to our use of a uniform electric field between the plates of an ideal parallel-plate capacitor in the earlier one.

equation 1



Magnetic field energy density

$$u_B = \frac{PE_B}{Vol}$$

$$u_B = \frac{B^2}{2\mu_0}$$

u_B = magnetic field energy density
PE_B = magnetic field energy
Vol = volume
B = magnetic field strength
μ_0 = permeability constant

Variables

We use Vol to represent volume in the definition of magnetic energy density to avoid confusion with the V in the symbol ΔV for electric potential difference.

magnetic potential energy in solenoid	PE_B
inductance of solenoid	L
current through solenoid	I
permeability constant	$\mu_0 = 4\pi\times10^{7}$ T·m/A
number density of loops	n
length of solenoid	s
area spanned by one loop	A
magnetic field strength	B
magnetic field energy density	u_B
volume inside solenoid	Vol

Strategy

1. Express the magnetic potential energy stored in the inductor (the solenoid) in terms of its inductance and the square of the current passing through it.
2. Replace the inductance and the current with expressions relating to the physical configuration of the solenoid, and the strength of the magnetic field inside it. This yields an equation for the potential energy.
3. State that the magnetic energy density equals the magnetic energy divided by the volume, and use the equation for potential energy just derived. Simplify to obtain the second equation on the right for magnetic field energy density.

Physics principles and equations

The magnetic potential energy stored in an inductor is

$$PE_B = \tfrac{1}{2} LI^2$$

The inductance of a solenoid is

$$L = n^2 s\mu_0 A$$

The formula for the magnetic field strength inside a solenoid is

$$B = \mu_0 In$$

The definition of magnetic field energy density is

$$u_B = \frac{PE_B}{Vol}$$

Step-by-step derivation

First, we derive a new expression for the energy PE_B stored in an inductor. For a solenoidal inductor, instead of relying on the inductance and current, the new expression will contain variables relating to the physical configuration of the solenoid and the magnetic field strength inside it.

Step	Reason
1. $PE_B = \tfrac{1}{2}LI^2$	energy of an inductor
2. $L = n^2 s\mu_0 A$	inductance of a solenoid
3. $I = \frac{B}{\mu_0 n}$	magnetic field strength of a solenoid
4. $PE_B = \tfrac{1}{2}(n^2 s\mu_0 A)\frac{B^2}{\mu_0^2 n^2}$	substitute equations 2 and 3 into equation 1
5. $PE_B = \frac{sAB^2}{2\mu_0}$	simplify

Now we state the definition of magnetic field energy density, and substitute the expression derived above for the magnetic potential energy inside the solenoid, along with an expression for the volume, into the definition.

Step	Reason
6. $u_B = \frac{PE_B}{Vol}$	definition of energy density
7. $Vol = sA$	volume of solenoid
8. $u_B = \frac{sAB^2/2\mu_0}{sA}$	substitute equations 5 and 7 into equation 6
9. $u_B = \frac{B^2}{2\mu_0}$	simplify

32.31 - Gotchas

Any magnetic field induces an emf in a loop of wire. No, a **changing** magnetic field induces an emf in such a loop. A uniform, constant magnetic field alone induces no net emf in a loop of wire. However, a potential difference can be induced across a segment of wire by moving it through a magnetic field.

If I double the strength of the magnetic field, I double the magnetic flux. Assuming nothing else has changed, you are correct.

If I double the surface area of the area enclosed by a wire loop in a magnetic field, I have doubled the magnetic flux. Again, assuming nothing else has changed, you are correct.

Transformers change potential differences. Power is a function of potential difference, so they change power, as well. No, the current changes in inverse proportion to the potential difference, so the power stays constant.

A battery and an inductor are in a series circuit. The emf of the inductor always opposes that of the battery. Not necessarily. If the current is increasing − say, a switch has just been closed − then this is true. But if the current is decreasing (perhaps the battery is abruptly failing), this change is also opposed, and the emfs have the same orientation. The inductor always acts to maintain the status quo.

32.32 - Summary

A changing magnetic field can generate an emf in a loop of wire. This phenomenon is called electromagnetic induction.

The motion of a coil in a nonuniform magnetic field can also induce an emf. The same effect is achieved whether you move the coil through the field, or move the field past the coil.

Magnetic flux, Φ_B is the amount of magnetic field passing perpendicularly through a surface. The amount of flux can be computed as the dot product of the field and area vectors, which is the product of the field strength, the area, and the cosine of the angle between them. The unit of magnetic flux is the weber (Wb). 1 Wb = 1 T·m2.

Faraday's law states that the induced emf in a circuit is proportional to the rate of change of magnetic flux through the circuit.

When a change in the magnetic flux due to an external magnetic field induces a current in a circuit, the induced current creates its own magnetic field. Lenz's law states that the induced magnetic field always opposes the change in flux that caused it. This helps you to deduce the direction of the induced current.

A loop of wire rotating in a magnetic field is the basis of an electric generator. A torque causes the loop to rotate in the field, constantly moving its wires through the field and inducing an emf and an electric current in it. The induced current reverses direction every half turn of the loop, and is called an alternating current (AC). Both the induced emf and the current vary sinusoidally with time.

Currents induced in large solid pieces of conducting material are not like those induced in a wire. When a solid conductor moves through a nonuniform magnetic field, eddy currents flow through it in complex "whirlpools." Eddy currents oppose the motion of the conductor through the nonuniform field.

Mutual induction occurs when a changing current in a circuit generates a changing magnetic field, which in turn induces a current in a second nearby circuit.

A transformer is a device that increases or decreases potential difference from one alternating-current circuit to another. Both circuits contain coils wrapped around the same iron core, but with a different number of loops in each coil.

Self-induction occurs when a changing current in a circuit causes an induced emf in the same circuit. The induced emf opposes the change in current that caused it, in accordance with Lenz' law. An inductor is a circuit component that exhibits substantial self-induction. A solenoid is a common configuration for an inductor.

Inductance is a property of inductors. It is equal to the negative of the induced emf divided by the rate of change in the current that causes it. The negative sign reflects the fact that the induced emf opposes the change that causes it. The symbol for inductance is L and it is measured in henrys (H). 1 H = 1 T·m^2/A.

An RL circuit contains a resistor and an inductor in series. If an emf source such as a battery is introduced into the circuit, the inductor causes current to increase more slowly than if the inductor were not present. An inductor stores energy in the magnetic field it creates.

Magnetic field energy density is the energy stored in a magnetic field per unit volume.

Equations

Wire segment moving in B field

$\Delta V = LvB$

$\mathcal{E} = LvB$

Magnetic flux

$\Phi_B = \mathbf{B} \cdot \mathbf{A}$

$\Phi_B = BA \cos \theta$

Faraday's law

$$\mathcal{E} = -N\frac{\Delta \Phi_B}{\Delta t}$$

"Induction-powered" circuit

$$P = \frac{L^2 v^2 B^2}{R}$$

Generator

$\mathcal{E} = NBA\omega \sin \omega t$

Transformer

$$\frac{\Delta V_1}{\Delta V_2} = \frac{N_1}{N_2} = \frac{I_2}{I_1}$$

Inductance

$$L = -\frac{\mathcal{E}}{\Delta I/\Delta t} \qquad L = N\frac{\Phi_B}{I}$$

$$\mathcal{E} = -L\frac{\Delta I}{\Delta t}$$

Chapter 32 Problems

Chapter Assumptions

Unless stated otherwise, use 5.00×10^{-5} T for the strength of the Earth's magnetic field at its surface.

Elementary charge, $e = 1.60\times10^{-19}$ C

Conceptual Problems

C.1 The textbook displays a photograph of an experiment where a bar magnet is dropped through a coil of wire. The changing magnetic flux through the coil induces an emf that in turn drives a current through the wire. The emf is displayed on an oscilloscope. After the magnet goes through the coil does it hit the pillow with the same kinetic energy as it would if there were no coil? Explain.

○ Yes ○ No

C.2 A large spool holding a coil of hundreds of loops of telephone cable is loaded on a truck at the factory in New York, and transported to a utility in Georgia. During the trip, a meter which is attached to the coil registers a tiny intermittent current in the coil. If there is no battery attached to the coil, what is the cause of the current?

C.3 A circular loop of wire is located in a constant, uniform magnetic field. How can an emf be induced in the loop?

C.4 A wire loop lies in the *xy* plane. A magnetic field points in the positive *z* direction. If the radius of the loop begins to shrink, in what direction will the current in the loop flow? Give your answer from the perspective of looking down on the *xy* plane, and explain your answer.

○ Clockwise ○ Counterclockwise

C.5 Suppose you are looking down on a wire loop, perpendicular to the plane it lies in. A magnetic field is pointed toward you and is increasing in strength. Will the induced current flow clockwise or counterclockwise? Explain your answer.

○ Clockwise ○ Counterclockwise

C.6 One solenoid is nested inside another of larger radius. Both solenoids are wired with the same "handedness." The larger solenoid has a current passing through it so that it generates a uniform interior magnetic field that surrounds the inner solenoid. The current in the outer solenoid begins to decrease, inducing a current in the inner solenoid. (a) In what direction does the current in the inner solenoid flow? Explain. (b) If the current in the outer solenoid remains steady, but the inner solenoid is drawn out of its open end, this process may induce a current in the inner solenoid. In what direction does the current in the inner solenoid flow? Explain.

(a) i. Same as outer solenoid
ii. There is no induced current
iii. Opposite to outer solenoid

(b) i. Same as outer solenoid
ii. There is no induced current
iii. Opposite to outer solenoid

C.7 "Superman - The Escape" is a thrill ride at Six Flags Magic Mountain in Valencia, CA. A 3000 kg car is accelerated to a height of over 100 meters. On the return trip the car is brought to a stop using electromagnets which induce eddy currents in the large wheels of the car. Since energy must be conserved, where does the nearly three million joules of gravitational potential energy go?

C.8 Carefully observe the animation of the swinging block in the textbook. (a) What is the direction of the eddy currents when the block swings into the magnetic field? (b) What is the direction of the eddy currents when the block swings out of the magnetic field? (c) Explain your answer to part a using Lenz's law.

(a) ○ Clockwise ○ Counterclockwise
(b) ○ Clockwise ○ Counterclockwise
(c)

C.9 The poles of a powerful magnet are oriented to oppose one another, with a gap in between them, so that the field lines from the north to the south pole are horizontal. You drop a copper penny so that it falls through the gap. (a) Describe what happens to the acceleration of the penny as it passes between the poles. (b) The effect described in part a would be different for a coin like a dime, which is close to the size and weight of a penny, but is made of a metal with a different resistivity. Describe a practical application of this difference in the coins, in a type of machine you have probabaly used.

(a)
(b)

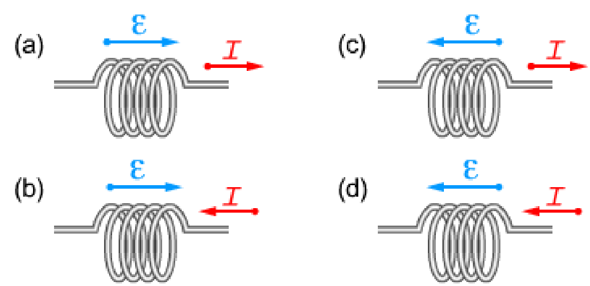
C.10 We have seen that it is possible for a changing current in one loop to cause a current to flow in an adjacent loop. (a) If the loops are lying flat on a surface, side by side, and the current in the first loop suddenly increases, does the induced current in the second loop flow in the same direction (by "direction" we mean clockwise or counterclockwise) as the original current, or in the opposite direction? Explain your answer. (b) Describe a configuration for the two loops which gives the opposite result.

(a) ○ The same direction ○ The opposite direction

(b)

C.11 For each of the four diagrams tell whether the current is increasing or decreasing based on the direction of the current flow and the direction of the induced emf.

(a) ○ Increasing ○ Decreasing
(b) ○ Increasing ○ Decreasing
(c) ○ Increasing ○ Decreasing
(d) ○ Increasing ○ Decreasing



C.12 A solenoidal inductor is made out of springy loops of wire. If the solenoid is stretched to twice its original length, what is the ratio of the original inductance to the new inductance?

____

C.13 The current through an inductor is increased by a factor of ten. By what factor, if any, does the magnetic potential energy stored in the inductor increase?

____

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Will a current flow in the bottom circuit if the current in the top circuit is steady and unchanging? (b) Is there current in the bottom circuit if the current in the top circuit is changing? (c) Does the rate at which the current in the top circuit changes have any effect on the potential difference you measure in the second circuit?

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c) ○ Yes ○ No

Section 2 - Interactive problem: motional induction

2.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Does the induced current change direction when the wire moves in a different direction? (b) What happens to the induced emf when you move the wire faster? (c) For a given wire speed, what happens to the induced emf when the strength of the magnetic field increases? (d) A controller in the simulation allows you to change the magnetic field strength, but only stepwise rather than continuously. Describe what happens to the induced emf when you impose a sudden jump in field strength.

(a) ○ Yes ○ No

(b) The induced emf
i. increases
ii. decreases
iii. stays the same
when the wire moves faster.

(c) The induced emf
i. increases
ii. decreases
iii. stays the same
when the field strength increases.

(d)

Section 5 - Motional induction: calculating the potential difference

5.1 A wire is moving perpendicularly through a magnetic field of strength 0.100 T with speed 60.0 m/s. If the potential difference between the ends of the wire is 3.00 V, what is the length of the wire?

__________ m

5.2 A sailboat with an aluminum mast 22.5 m high is sailing eastward along the equator at a speed of 3.66 meters per second. The Earth's magnetic field at the ship's location is 5.50×10^{-5} T, horizontal and directed due north. What is the induced potential difference between the top and bottom of the mast?

________ V

5.3 A Boeing 777-200 airliner with a wingspan of 64.0 m is flying perpendicular to the Earth's magnetic field at its cruising speed of 250 m/s. The strength of the field is 5.05×10^{-5} T. What potential difference is induced between the plane's wingtips?

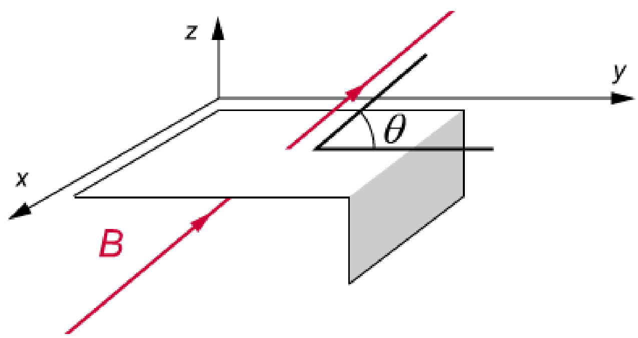
________ V

5.4 During your biking tour of the Pacific Northwest, you decide to outdo those latte-guzzling ecotopians at their own game. Instead of the four toxic and disposable 1.5-volt batteries suggested for your walkabout CD player, you will place a 2.0 m conducting wand across your handlebars. Pedaling furiously in a direction perpendicular to the local magnetic field, you wish to go fast enough to induce a potential difference of 6.0 V (equivalent to four batteries in series). (a) The strength of the Earth's magnetic field at your location is 5.0×10^{-5} T. How fast will you need to go? (b) Is this speed street-legal? (c) The magnetic declination at your location is 19°. At what angle from geographic north should you travel to maximize the potential difference of your environment-friendly arrangement?

(a) ________ m/s
(b) ○ Yes ○ No
(c) i. 0 degrees
ii. 19
iii. 71
iv. 90

Section 6 - Magnetic flux

6.1 A 0.150 T magnetic field *B* crosses the *xy* plane perpendicular to the *x* axis and at an angle of 30.0° to the *y* axis. A square surface with sides of length 7.00 m lies mainly in the *xy* plane, but a 1.50 m section of it at one end is bent 90° in the negative *z* direction. What is the total magnetic flux through the surface? Choose directions for the area vectors so that the flux through each surface is positive.



________ Wb

6.2 A wire circle of radius 0.050 m is embedded in a solenoid of length 0.20 m with 1000 turns that carries a current of 0.50 A. If a vector that is normal to the plane of the circle makes a 40° angle with the axis of the solenoid, what is the magnetic flux through the loop?

________ Wb

Section 7 - Faraday's law

7.1 A magnetic field increases steadily in strength, starting at 0 T. The induced emf in a loop that is resting in the field is 2.72 V. After 4.75 s, what is the magnitude of the magnetic flux through the loop?

________ $T \cdot m^2$

7.2 A coil with multiple loops is in a magnetic field such that the magnetic flux through the coil increases at a rate of 0.0240 Tm^2/s. The induced emf in the coil is −3.36 V. How many loops are in the coil?

________ loops

7.3 A wire has a resistance of 60.0 ohms, and is bent into a square loop with sides of length 15.0 cm. The loop is perpendicular to a magnetic field which is changing at a rate of 3.00e−3 T/s. What is the current induced in the wire?

________ A

7.4 You hold a thin gold wedding ring with an inside radius of 1 cm, such that the plane of the ring is perpendicular to the Earth's magnetic field. At your location it has strength 5.5×10^{-5} T. What is the emf induced in the ring if you turn it in one-tenth of a second to be parallel to the field (that is, you make the area vector of the ring perpendicular to the field)?

________ V

Section 8 - Interactive problem: Faraday's law

8.1 Use the information given in the interactive problem in this section to calculate the time interval required to induce the desired emf in the coil. Give your answer in milliseconds. Test your answer using the simulation.

_________ ms

Section 9 - Flux and motional induction

9.1 A U-shaped wire track has a 144 Ω light bulb wired into it at the base of the "U". A straight wire segment of length 5.46 cm is moved along the parallel sides of the "U", away from the bulb, and perpendicular to a surrounding uniform magnetic field. This generates an emf in the circuit and causes the bulb to glow. (a) If the wire moves at a speed of 0.895 m/s, and the strength of the magnetic field is 45.6 mT, what is the power consumed by the bulb? Assume for this part of the problem that the wires in the apparatus have no resistance. (b) Suppose the speed is doubled. What happens to the power consumption? (c) Suppose that the wires in the apparatus have a small but nonzero resistance, and that the parallel sides of the "U" are infinitely long. What happens to the brightness of the bulb over a long period of time as the wire slider moves away at a constant velocity?

(a) _________ W

(b) The power i. increases by a factor of _________
ii. decreases

(c)

9.2 In Chris Van Allsburg's story *The Polar Express*, a boy takes a magical train ride to the North Pole. Near the pole, assume the strength of the Earth's magnetic field is 5.02×10^{-5} T and the field lines are vertical (as they actually are at the nearby magnetic south pole). Also assume that at the end of the track, the two steel rails are connected by a conducting wire. Each axled pair of steel wheels on the train closes a circuit with the end of the track. The width of the track is the standard US/Canadian gauge, 1.44 m, and the train is moving at 26.0 m/s. (a) What emf is generated in the rails ahead of the train? (b) In part a, you are not told how many pairs of wheels the train has. Why don't you need to know?

(a) _________ V

(b)

Section 13 - Physics at work: an "induction-powered" flashlight

13.1 A circuit consists of two very low-resistance rails separated by 0.300 m, with a 50.0 ohm resistor connected across them at one end and a conducting, moveable bar at the other end. The circuit is in a uniform 0.100 T magnetic field that is parallel to the area vector of the circuit. The bar is moving at a constant velocity such that the resistor dissipates 1.50 W of power. What is the speed of the bar?

_________ m/s

13.2 A pair of parallel conducting rails is separated by 55.0 cm. They are joined at one end by a circuit component containing a light bulb with resistance 144 Ω. A transverse rod completes the circuit, sliding down the rails away from the bulb at a speed of 3.25 cm/s. Finally, the whole apparatus is located in a uniform magnetic field perpendicular to the plane of the circuit, and having a strength of 7.44 T. There is an induced emf in the circuit causing the bulb to glow. How much power is being consumed by the bulb?

_________ W

13.3 Electrodynamic space tethers are described in the section Interactive checkpoint: space tether. A particular space tether is a wire made from aluminum (resistivity 2.7×10^{-8} Ωm, density 2700 kg/m^3). Its mass is 10.0 kg and it moves perpendicularly through a uniform 20.0 μT magnetic field at a speed of 6.81×10^3 m/s. (a) What power is produced by the tether's motion through the magnetic field? Hint: Use the definition of resistance in terms of resistivity and the definition of density. (b) If the tether is 2.00 km long with a uniform cross-section, what is the area of the cross-section? (c) What is the overall resistance of the 2.00 km long tether?

(a) _________ W

(b) _________ m^2

(c) _________ Ω

Section 15 - Interactive problem: a generator

15.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) With what orientation of the velocity vector is the emf a maximum? (b) With what orientation of the velocity vector is the emf zero? (c) Does the frequency of the emf change with angular velocity? (d) Does the amplitude of the emf change with angular velocity?

(a) The emf is a maximum when the velocity vector is i. Perpendicular ii. Parallel to the magnetic field.

(b) The emf is zero when the velocity vector is i. Perpendicular ii. Parallel to the magnetic field.

(c) ○ Yes ○ No

(d) ○ Yes ○ No

Section 16 - Quantifying the emf induced in a generator

16.1 A generator is made from a set of 50 wire loops of radius 40.0 cm that turn in a uniform 0.0900 T magnetic field. What angular frequency of rotation is needed to supply a peak emf of 85.0 V?

__________ rad/s

16.2 The peak emf of a generator is 30.0 V when the armature of the generator turns at 1200 rpm (revolutions per minute). What is the peak emf when the rate of rotation is increased to 4600 rpm?

__________ V

16.3 A generator is made from a set of 7000 loops of area 5.00 cm^2 rotating at an angular frequency of 600 radians per second. The magnetic field is created by a solenoid, completely surrounding the generator coil, and carrying a current of 0.300 A. How many turns per unit length must the solenoid have in order for the generator to generate a peak emf of 1000 V?

__________ turns/m

Section 17 - Interactive problem: configuring a generator

17.1 Use the information given in the interactive problem in this section to answer the following question. If the number of loops in the generator coil is equal to 4, the magnitude of the magnetic field is 5.00 T, and the angular velocity is 4.70 rad/s, what is the loop area required to induce the desired emf? State your answer to the nearest 0.10 m^2 and test it using the simulation.

__________ m^2

17.2 Use the information given in the interactive problem in this section to answer the following question. If the loop area is equal to 2.00 m^2, the magnitude of the magnetic field is 6.00 T, and the angular velocity is 3.50 rad/s, what is the number of loops required to induce the desired emf? Test your answer using the simulation.

____

Section 20 - Transformers

20.1 A transformer with an input current of 0.500 A has 6000 primary turns. How many secondary turns must the transformer have if an output current of 2.00 A is desired?

__________ turns

20.2 A machine that makes transformers can only wind the coils with a single specific number density. The number of turns is therefore controlled by the length of the coils. A transformer is designed to transform a 120 V input potential difference to a 9.0 V output. If the input coil is 2.8 cm in length, what is length of the output coil?

__________ m

20.3 The long distance power transmission lines reaching a certain power substation operate at 375,000 V. This potential difference has to be "stepped down" for local distribution through city power lines at 14,300 V. Substations use very large, oil-cooled transformers (a photograph of one is in the text) to accomplish this conversion. What is the ratio of primary to secondary windings in a transformer in such a substation?

____

20.4 The electricity supplied through a city power line is transmitted at 14,300 V. On each city block there are one or more neighborhood transformers on utility poles that transform the potential difference to 120 V, as used in homes and businesses (240 V outlets are supplied by two 120 V lines). What is the ratio of primary to secondary windings in a neighborhood transformer like the one shown in the photograph?

______

20.5 (a) The emf produced by a typical hydroelectric generator, at the generator, is 20,000 V. The electricity produced by the generator is fed through a step-up transformer before being transmitted across the country on "high tension" lines like those shown in the textbook. If the primary winding of the transformer has 1550 loops, what must be the multiplicity of the secondary winding to produce a transmission emf of 100,000 V? Of 225,000 V? Of 345,000 V? (b) A current of 4500 A flows from the generator to the step-up transformer. What current flows through the transmission lines at each one of the three operating emfs stated in part a? (c) If the total resistance of the power lines between the generator and a city it supplies with electricity is 112 Ω, use the results of part b to calculate the power dissipated in the lines at each of the three operating emfs. (d) In commercial or public power systems, why is electric power transmitted at extremely high emfs and (relatively) low currents, rather than the other way around?

(a) ______ for 100,000 V ______ for 225,000 V ______ for 345,000 V
(b) ______ A for 100,000 V ______ A for 225,000 V ______ A for 345,000 V
(c) ______ W for 100,000 V ______ W for 225,000 V ______ W for 345,000 V
(d)

Section 22 - Interactive problem: configuring a transformer

22.1 Use the information given in the interactive problem in this section to determine the maximum potential difference in the primary coil. Test your answer using the simulation.

______ V

Section 24 - Inductors and inductance

24.1 A 350 mH solenoidal inductor has an induced emf of 12 V across it when the current flowing through it is changing at a certain rate. What is the rate of change of the current? Report your answer as a signed number.

______ A/s

24.2 What is the inductance of a solenoid that produces an induced emf of −3.4 V when the current flowing through it is changing at the rate of 0.15 A/s?

______ H

24.3 A scientist measures the current flowing through a 2.35 H inductor and finds that it is 0.822 A. 5.00 minutes later she checks the current and finds that it has decreased to 0.545 A. She knows that in her experimental apparatus the current varies in a linear fashion. What emf was induced across the inductor during the five-minute period?

______ V

24.4 A certain solenoidal inductor has a magnetic flux of 835 μWb through each loop of its coil when the current flowing through it equals 0.600 A. If the solenoid has 125 loops, what is its inductance?

______ H

Section 27 - Calculating the inductance of a solenoid

27.1 A metal Slinky® toy, with 64 loops and a diameter of 7.0 cm, is stretched to a length of 0.95 m. What is its inductance?

______ H

27.2 An inductor is to be made by wrapping wires around a paper tube of radius 0.702 cm and length 8.44 cm. How many turns will give the inductor an inductance of 3.00e−3 H?

______ turns

Section 29 - Energy stored in an inductor

29.1 What current must be passing through a 0.0700 H inductor so that it stores 0.300 J of energy?

______ A

29.2 A 5.00 mH inductor has a current of 7.50 mA flowing through it, while a 5.00 mF capacitor holds a charge of 7.50 mC. (a) What is the magnetic potential energy stored in the inductor? (b) What is the electric potential energy stored in the capacitor? (c) The inductance and the capacitance are both tripled. What is the ratio of the new potential energy in the inductor to the old, and of the new potential energy in the capacitor to the old? (d) The inductance and capacitance are returned to their original values, and the current and the capacitor's stored charge are both doubled. Now what are the ratios of the new potential energies to the old?

(a) __________ J
(b) __________ J
(c) Inductor:

i. 4
ii. 3
iii. 2
iv. 1/2
v. 1/3
vi. 1/4

Capacitor:

i. 4
ii. 3
iii. 2
iv. 1
v. 1/2
vi. 1/3
vii. 1/4

(d) Inductor:

i. 4
ii. 3
iii. 2
iv. 1/2
v. 1/3
vi. 1/4

Capacitor:

i. 4
ii. 3
iii. 2
iv. 1/2
v. 1/3
vi. 1/4

Section 30 - Magnetic field energy density

30.1 A magnetic field and an electric field, in separate regions of space, have strengths of 1.25 T and 1.25 N/C, respectively. (a) What is the energy density of the magnetic field? (b) What is the energy density of the electric field? (c) If the magnetic and electric field strengths are doubled, in each case what is the ratio of the new energy density to the old one? (The answer is the same whether you consider the magnetic or the electric field.)

(a) __________ J/m^3
(b) __________ J/m^3
(c) __________

30.2 A solenoid is made from wire, and has length 0.030 m, radius 5.0×10^{-3} m, and 1500 total turns. A current of 0.25 A runs through it. What is the magnetic energy density inside the solenoid?

__________ J/m^3

Chapter 33 Alternating Current Circuits

33.0 - Introduction

This radio employs an *RLC* circuit to tune in various radio frequencies.

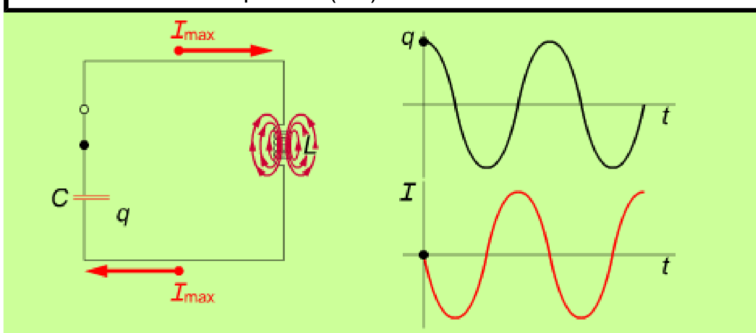
When you studied direct current circuits, you concentrated on analyzing circuits that were in a steady state. The potential differences and currents in most of the circuits you analyzed remained constant over time; time does not appear as a variable in equations such as Ohm's law or Kirchhoff's rules.

However, for an important class of circuits, timing is everything. In alternating current (AC) circuits, both the current passing through, and the potential differences across components are constantly changing.

One type of AC circuit is called an *RLC* circuit, after properties that define its components: resistance (*R*), inductance (*L*), and capacitance (*C*). One use of this type of circuit is to "tune in" radio signals of different frequencies. If you turn the dial of an old-style radio, you may be changing the capacitance of a capacitor. This change in capacitance alters the timing − what is called the resonance frequency − of the radio's receiver circuit to match the broadcast frequency.

AC circuits constitute a complex subject, which we will explore step by step. First, we explain how the components function in a circuit without an AC (alternating current) generator. Then, we examine separately the properties of each component when it is placed in a circuit that contains an AC generator. Finally, we discuss what occurs when all the components are placed together in a series circuit powered by a generator.

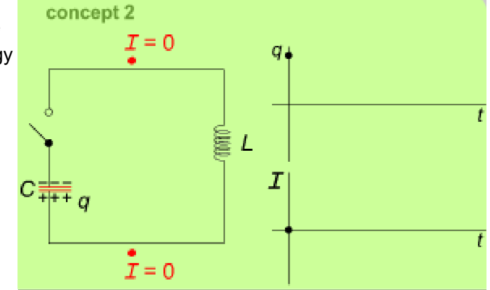
33.1 - Inductor-capacitor (LC) circuits



concept 1

Inductor-capacitor (*LC*) circuits
Current flows back and forth in circuit
Capacitor, inductor alternate "driving" current

concept 2



Capacitor fully charged, no current

In this section, we analyze the ebb and flow of the current and other properties of an inductor-capacitor (*LC*) series circuit. Your understanding of the functioning of this type of circuit will be aided by recalling some useful relationships. First, the amount of energy stored by a capacitor is proportional to the square of the charge on each plate. The more charge you see on a capacitor plate, the more energy it is storing.

Second, the amount of energy stored in an inductor is proportional to the square of the current flowing through it. If you note how much current is flowing through the inductor, you have a sense of how much energy is stored in its magnetic field.

In this section, we discuss a circuit without a generator. Initially, the switch is open and the capacitor plates are charged. To simplify the analysis, we ignore the very small resistance present in the wires of the circuit, which means that energy is conserved within the circuit, and none is dissipated as heat.

The current, the charge on the capacitor plates, and the magnetic field of the inductor will continuously change once the switch is closed. We will now walk through this process step by step, starting at the time just before the switch is closed. Each of these steps is shown on the right. You can also see the circuit functioning dynamically in the animation above.

1. **The capacitor is fully charged.** There is a potential difference across the capacitor because it is fully charged. All the energy in the circuit is now stored here. Since the switch is open, no current flows. No current means the inductor does not have a magnetic field.
2. **The switch is closed. The capacitor begins to discharge, creating a current.** With the switch closed, the potential difference across the capacitor causes a current to flow and the capacitor begins to discharge. The potential difference across the capacitor decreases as it discharges. The current through the inductor creates a magnetic field. Remember: The inductor's magnetic field is proportional to the

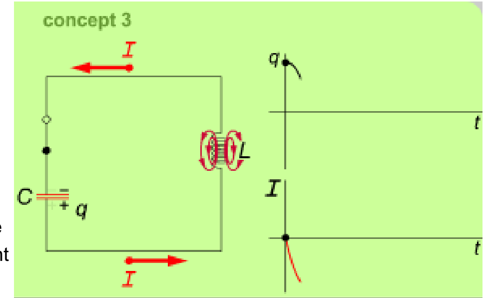
amount of current flowing through it.

3. **The current is at its peak amount, the inductor's magnetic field is at its maximum, and the capacitor is fully discharged.** At this instant, the capacitor plates have no charge, which means there is no potential difference across the capacitor. All the energy in the circuit is now stored in the magnetic field of the inductor. This means the current is at its maximum amount, since the energy in the inductor is a function of the square of the current.
4. **The current continues, driven by the induced emf caused by the reduction of the magnetic field in the inductor, and recharges the capacitor.** Since the capacitor is fully discharged, there is no potential difference across it. The current diminishes, which reduces the magnetic field of the inductor. This change in magnetic flux induces an emf that opposes its cause, the reduction in current. The induced emf causes the current to continue to flow in the same direction as before. This recharges the capacitor plates.
5. **The current ceases as the inductor exhausts the energy stored in its magnetic field; the capacitor again reaches its peak charge.** The circuit returns to its initial condition, except that the capacitor's charges are now reversed (the plate that was positive is now negative and vice versa). The magnitude of the potential difference across the capacitor returns to its initial value.
6. **The process repeats but with the current flowing in the opposite direction.** Once again, the capacitor discharges. The current now flows in the opposite direction. As the current flows, the capacitor discharges and the inductor once again possesses a magnetic field (which now points in the opposite direction than it did in step 3). When the capacitor is fully discharged, the inductor once again drives the current, in the process recharging the capacitor to its initial state.

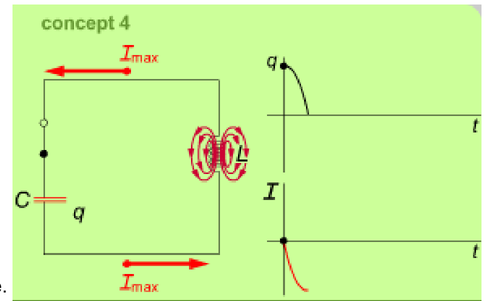
One cycle has now been completed. If you look at Concept 7, you will see the circuit has returned to its initial state as shown in Concept 1 (although the switch is now closed). There is no current at this instant, and the capacitor is charged as before. In principle, the process described above in this idealized circuit will repeat forever.

The analysis you have just followed introduces some of the basic themes of this chapter. First, the circuit can be analyzed in terms of energy. You can think of an *LC* circuit as one in which energy flows from the capacitor to the inductor and then back to the capacitor.

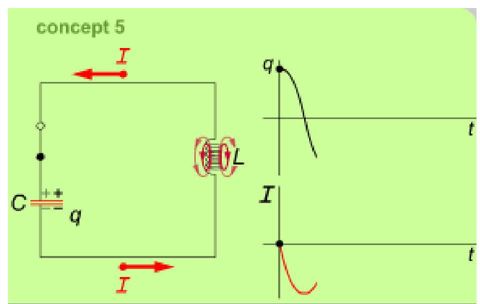
Second, time figures prominently in this kind of circuit. The current and charge vary over time, and there is also a consistent relationship between them. You can see this in the graphs on the right. The charge on the capacitor is at its maximum magnitude when the current is zero, and the converse is true as well. When there is no charge on the capacitor, the current is at its maximum magnitude.



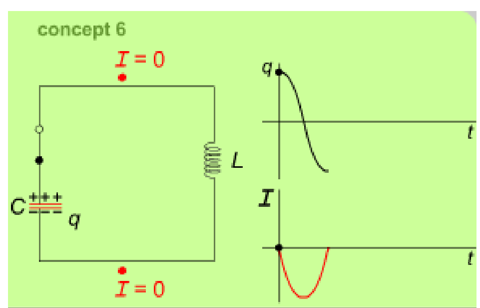
Capacitor discharges



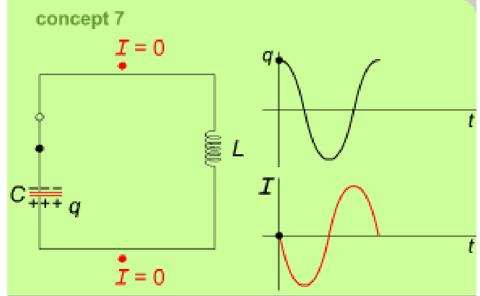
Capacitor fully discharged



Inductor drives current

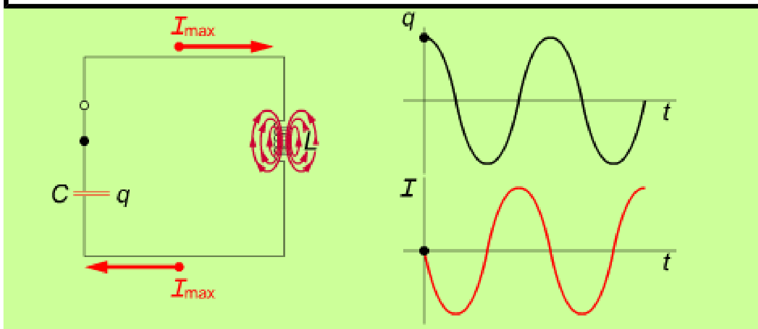


Capacitor fully charged, polarity reversed



Process continues in reverse direction
Circuit returns to original state

33.2 - LC oscillations quantified



concept 1

Quantifying *LC* oscillations
Charge and current vary sinusoidally

Equations can be used to describe the capacitor charge and the current in an *LC* circuit as a function of time.

The equations in Equation 1 and Equation 2 use sinusoidal functions to describe the charge on a capacitor plate and the current in the circuit. The equation for the charge on a capacitor plate includes the factor q_{max}, the maximum charge. The argument of the sine function, $\omega t + \pi/2$, includes the phase constant $\pi/2$ (90°), which is used to account for the initial fully charged state of the capacitor. When t is zero, $\sin(\omega t + \pi/2)$ becomes $\sin \pi/2$, which equals one, so that q equals q_{max}.

In Equation 2 you see the equation for the current. It equals the negative of the product of the maximum current I_{max} and $\sin \omega t$.

Notice that the phase constant in this equation for the current is zero, while the phase constant in the equation for the charge is $\pi/2$. This accounts for the relationship in time between the current and the charge. Initially, the charge is at its maximum and there is no current. Conversely, when the charge is zero, the current is flowing at its maximum rate in one direction or the other.

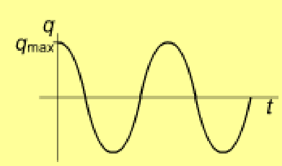
You can also account for the timing relationship between charge and current by considering their mathematical relationship. The current equals the rate of change of the charge on the capacitor plates, which is the slope of the charge function shown in the graphs in this section. When a line tangent to the curve is the steepest, the slope is at its greatest. For instance, when the capacitor is discharging at its maximum rate, the current is at its greatest magnitude.

If you look at the graphs atop this page, you can confirm that the magnitude of the slope of the charge graph equals the magnitude of the current graph below it. When the charge is at its maximum, the slope is zero (since a line tangent to the curve is horizontal), and the current equals zero. Conversely, when the charge is zero, the slope of the charge curve is at its positive or negative maximum, as is the magnitude of the current.

The formula for the **maximum** current in this circuit is also shown in Equation 2. It equals the angular frequency of the circuit times the maximum charge. These equations can be derived using differential equations and the law of conservation of energy.

The angular frequency, ω, measured in radians per second, is a function of the capacitance of the capacitor and the inductance of the inductor, as shown in Equation 3. Increasing either value reduces the rate of oscillation. (For a quick review: Angular frequency equals 2π times the frequency. Frequency is measured in cycles per second, or hertz. The period is the time required for one cycle, and equals the reciprocal of frequency.)

equation 1

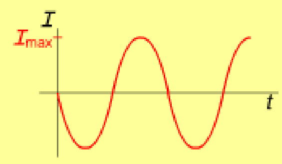


Charge on capacitor

$$q = q_{max} \sin(\omega t + \pi/2)$$

q = charge on capacitor plate
q_{max} = maximum charge on plate
ω = angular frequency
t = time
$+\pi/2$ = phase constant

equation 2



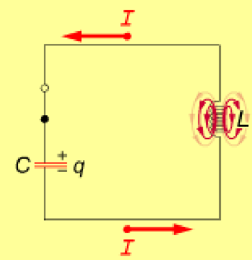
Current in circuit

$$I = -I_{max} \sin \omega t$$

$$I_{max} = \omega q_{max}$$

I = current
I_{max} = maximum current

equation 3

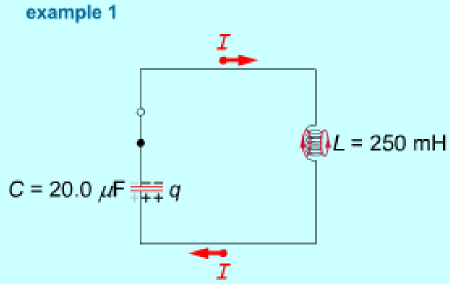


Angular frequency of oscillations

$$\omega = \frac{1}{\sqrt{LC}}$$

L = inductance
C = capacitance

example 1



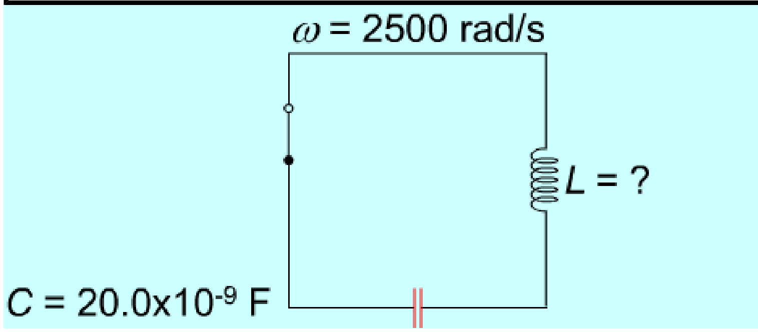
What is the angular frequency of the oscillations in this circuit?

$$\omega = \frac{1}{\sqrt{LC}}$$

$$\omega = \frac{1}{\sqrt{(250 \times 10^{-3}\text{H})(20.0 \times 10^{-6}\text{F})}}$$

$$\omega = 447 \text{ rad/s}$$

33.3 - Sample problem: LC angular frequency



An LC circuit with a 20.0x10^{-9} F capacitor needs to have an angular frequency of 2500 rad/s for a certain application. What inductance is needed in the circuit?

Variables

capacitance	$C = 20.0\times10^{-9}$ F
angular frequency	$\omega = 2500$ rad/s
inductance	L

What is the strategy?

1. Use the equation for the angular frequency of an *LC* circuit.
2. Solve for the needed inductance.
3. Substitute known quantities and evaluate.

Physics principles and equations

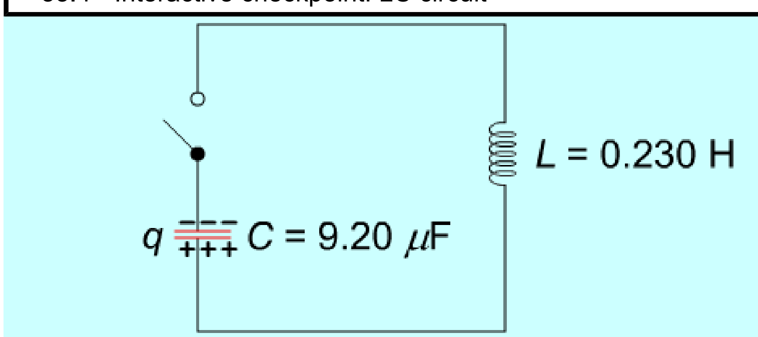
The angular frequency of an *LC* circuit is

$$\omega = \frac{1}{\sqrt{LC}}$$

Step-by-step solution

Step	Reason
1. $\omega = \dfrac{1}{\sqrt{LC}}$	angular frequency of LC circuit
2. $L = \dfrac{1}{C\omega^2}$	solve for L
3. $L = \dfrac{1}{(20.0\times10^{-9}\text{F})(2500\text{ rad/s})^2}$ $L = 8.00$ H	evaluate

33.4 - Interactive checkpoint: LC circuit



A 0.230 H inductor is wired in series with a 9.20 μF capacitor. The capacitor is fully charged when the circuit is closed at time t = 0. At what time will the charge fall to half its initial value? State the first time at which this occurs.

Answer:

t = ______ s

33.5 - An AC generator, its emf, and AC current

In this section, we consider the implications of adding an *AC* (*alternating current*) generator to a circuit. We start with the circuit you see on the right, one that contains only a generator and a resistor. The AC generator is represented by a blue circle with a wave inside.

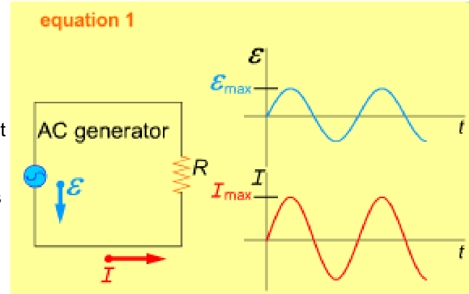
At its simplest, an AC generator consists of a wire loop that rotates in a magnetic field. It generates an emf that varies sinusoidally with time.

The emf created by the generator can be described mathematically as the product of its maximum emf and a sine function, as shown in Equation 1. The factor ω in the argument of the function is the angular frequency of the current. In the United States and East Asia, alternating current has a **frequency** of 60 cycles per second (Hz), meaning the emf reaches its maximum positive value 60 times per second. The **angular frequency** is 2π times this value, about 377 rad/s.

In the United States, the maximum emf of alternating current is approximately 170 volts. Utilities strive to provide a potential difference that, calculated in a specific way, averages to 120 volts.

When an AC generator is placed in a circuit, its emf causes a current to flow. We can now use principles applied earlier to direct current circuits to analyze this circuit.

First, Kirchhoff's loop rule states that the sum of the potential differences as a closed loop is traversed equals zero. This means that the emf of the generator is equal in magnitude to the potential difference across the resistor. Second, using Ohm's law, we can divide the potential difference by the resistor's resistance to determine the current. Since the emf and potential difference vary sinusoidally, so does the current. The equation for the current is also shown in Equation 1. We will further explore the relationships between current and emf in an AC circuit shortly.



AC generator

$$\mathcal{E} = \mathcal{E}_{max} \sin \omega t$$

$$I_R = I_{max} \sin \omega t$$

$\mathcal{E}$ = emf
I_R = current through resistor circuit
ω = angular frequency
t = time

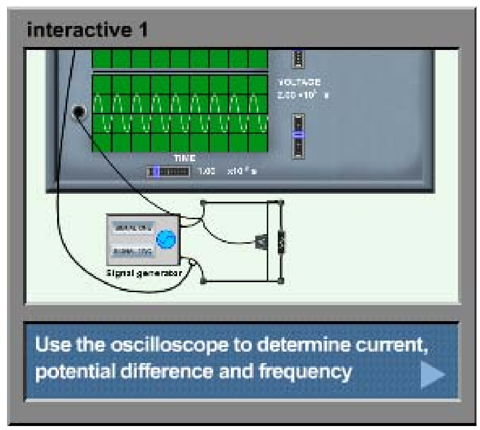
33.6 - Interactive problem: alternating current

In this simulation, there is a circuit that contains a signal generator and a resistor. The emf of the signal generator causes a current to flow. Your task is to use an oscilloscope to determine the maximum current I_{max}, the maximum potential difference across the resistor ΔV_{max} and the signal frequency f.

You will see sinusoidal graphs in the oscilloscope similar to those you see in illustrations in other sections of this chapter.

One probe from the oscilloscope is used to measure current. The current is the same everywhere in the circuit. The two ends of the potential difference ("voltage") probe in this exercise are placed so it measures ΔV across the resistor.

The oscilloscope displays information in the form of two graphs. In the upper graph, the current is displayed on the vertical axis and time is on the horizontal axis. In the lower graph, ΔV is displayed on the vertical axis, and time once again is on the horizontal axis.



Sliders allow you to change the scales of the graphs. The base unit is a single gray square in the oscilloscope window. For example, if you set the current slider to "8.00×10^{-2} A," then the height of a gray box represents 8.00×10^{-2} amperes. You can set the scale for the time in an analogous fashion. All this resembles the working of typical laboratory oscilloscopes.

You set the scale by trial and error. If you cannot see any signal, move the slider so the units are smaller until you can see the graph. If the waves you see are so large that you cannot see their peaks, increase the scale until their peaks "shrink" into view.

When you launch the simulation, the signal generator is on and all the probes are attached. The peak of the wave in the current display is the maximum current, I_{max}. The peak of the wave in the voltage display is ΔV_{max}.

You also have to analyze the wave to determine the frequency of the signal. To do so, identify two adjacent matching points in the wave (such as two peaks) in order to identify a cycle. The time it takes to complete this cycle is the period of the wave, and the reciprocal of the period is frequency. For instance, if the time between two peaks is 1.0×10^{-3} seconds, then the frequency is 1/(1.0×10^{-3}), or 1,000 hertz (Hz).

A couple of tips may help here. Since the horizontal axis measures time, you can determine the period by counting how many gray boxes there are between the corresponding points on the wave (be careful to note what the time scale is). It can be easier to use locations where the wave crosses the horizontal axis, since more precise values can be noted there. But if you use those points, remember that matching points are where the graph is in both cases going down, or in both cases going up. Two adjacent points where the wave crosses the horizontal axis, but in opposite directions, represent only half a cycle.

When you have analyzed a signal and determined I_{max}, ΔV_{max} and the signal frequency, enter these values in the boxes in the simulation. Press CHECK to see if you are correct.

To measure the second signal, press the Signal Two button on the signal generator. The signal generator will create a signal with a different potential difference and frequency. You can also analyze this signal, enter your conclusions and check them by pressing CHECK.

33.7 - Phasors

Phasors (spelled differently than Star Trek™ phasers, and only slightly less dangerous) are used to represent sinusoidal functions. They are particularly useful in showing the relationship over time of various quantities (such as charge, current and potential difference) that can be described by these functions.

A phasor is a vector that rotates through different angular positions. If you recall how to draw vectors using polar coordinates − length and angle − then you are well along the way to understanding phasors.

The same vector that can be specified in rectangular coordinates as (2.0, 2.0) can also be specified in polar coordinates as (2.8, $\pi/4$ radians). The first value, 2.8, represents the length (magnitude) of the vector, and $\pi/4$ radians $(45°)$ is the angle the vector makes with the positive x axis.

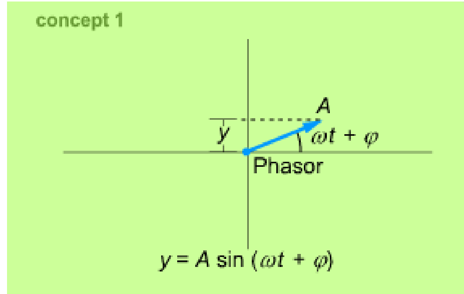
Phasors are useful in displaying expressions like the one shown in the illustration to the right, $A \sin(\omega t + \varphi)$. The length of the phasor is the amplitude A of the sinusoidal function. Its angle is the argument of the sine, $\omega t + \varphi$. The y component of the phasor (its "height", so to speak) is the value of the function at the time t. The angle changes with time as the phasor rotates, causing the y component to change as well. Only the length of the phasor remains constant.

Let's pick a specific function, $4 \sin(\pi t + \pi/6)$, and evaluate it at $t = 0$ to find its initial value. The angle at that time is $\pi/6$ rad $(30°)$. The length of the phasor equals the amplitude of the function, which is 4. At $t = 0$ the phasor is a vector with a length of four and an angle of $\pi/6$ rad. In polar notation, this is $(4, \pi/6)$. Its y component is the initial value of the function, 4 times the sine of $\pi/6$ radians, which equals 2.

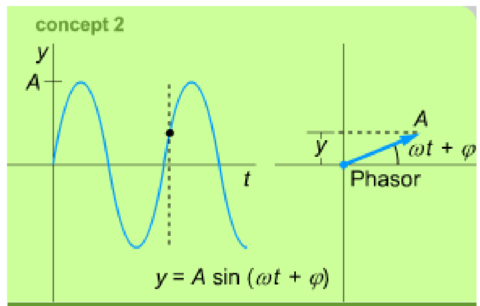
You may have plotted such a function as a sine wave when analyzing a wave or a particle moving in simple harmonic motion. You would have calculated the values of the expression at various times t and plotted them. At $t = 0$, the function equals 2, as mentioned above. Since the horizontal axis measures time, you would have plotted the point (0, 2).

Let's consider the value of the function at 1/3 second. The angle is $\pi/3 + \pi/6 = \pi/2$. On a graph, you would plot the point (1/3, 4). If you were using phasors, you would draw a vertical vector (angle = $\pi/2 = 90°$) of length 4.

Concept 2 contains an animation showing the same function displayed in the two fashions discussed in this section. The animation ends by showing the phasor at a particular instant in time. Click on Concept 2 to see the animation.



Phasors
Vector-based way to represent functions
A phasor is a vector with
· length equal to amplitude of function
· angle equal to argument $\omega t + \varphi$
· height equal to value of function



Phasor rotation
Phasor rotates as angle changes

The phasor will complete one revolution when its argument changes by 2π radians $(360°)$. How soon it accomplishes this depends on the angular frequency ω of the function. For instance, when $\omega = \pi$, as with the function $4 \sin(\pi t + \pi/6)$, the phasor will complete one revolution in two seconds.

33.8 - Phasors and phase differences

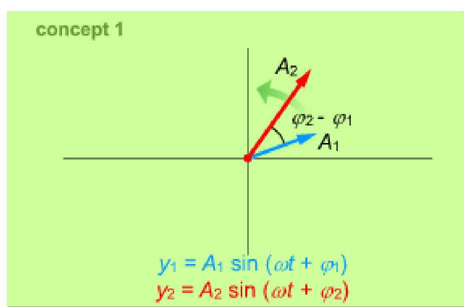
Phasors prove quite useful when two sinusoidal functions have the same angular frequency but different phase constants.

In the first illustration, you see two functions and the phasors that represent them. The functions have the same angular frequency, but they differ in their amplitudes and by their phase constants, φ_1 and φ_2. The functions have a phase difference of $\varphi_2 - \varphi_1$. The angle between the two phasors will always equal this difference between their phase constants.

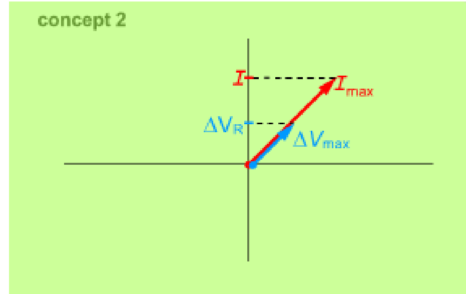
This example illustrates one useful aspect of phasors: It is easier to see the phase difference between these two functions in a phasor diagram than by comparing their graphs as waves.

The two phasors in Concept 2 have no phase difference. They are said to be "in phase" with each other. One represents the potential difference across a resistor in a circuit consisting of just the resistor and an AC generator, and the other represents the current in this circuit.

Phasors such as those in Concept 2 cannot be added as vectors. You cannot add potential difference and current any more than you can add apples and oranges. The vertical axis has two scales, one for potential difference and a separate one for current. (Some physicists do not like this shared-axis approach, and draw two distinct phasor diagrams to avoid potential confusion.)



Multiple phasors with same angular frequency
Angle between is phase difference $\varphi_2 - \varphi_1$
Angle between is constant during rotation



Parallel phasors
Have zero phase difference
· In example, ΔV and I are "in phase"

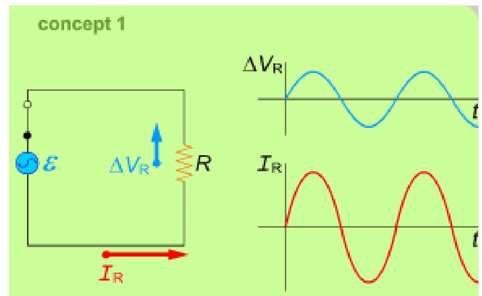
33.9 - Alternating current and a resistor

In this section, we analyze a circuit that consists of an alternating current generator and a resistor. This begins our discussion of the timing relationships among the various components of an *RLC* circuit.

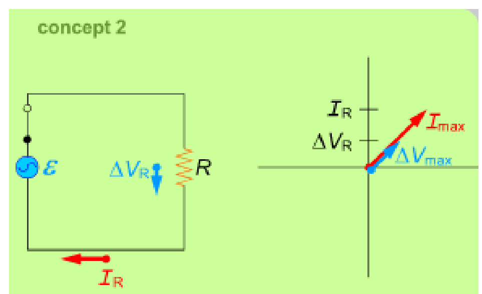
We use some of the essentials that were discussed in direct current circuits to analyze a circuit with an alternating current. Since there are just two components, we can use Kirchhoff's loop rule to conclude that the magnitude of the potential difference across the resistor equals the emf of the generator at all times. This means we can describe it with a sine function, just as we describe the emf of the generator.

The emf of the generator can be described with the equation $\mathcal{E} = \mathcal{E}_{\text{max}} \sin \omega t$. The potential difference across the resistor oscillates in phase with the emf. Ohm's law dictates that the current is proportional to the potential difference across the resistor. In terms of timing, this means that the current and the potential difference across the resistor are in phase. They peak and trough at the same times, and both are zero at the same times as well. In terms of equations, it means that each can be described as the product of its amplitude and the function sin (ωt).

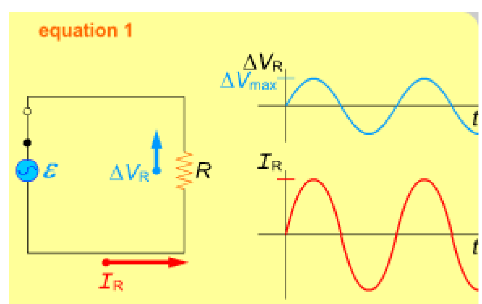
You see both the equations and the phase relationships on the right. Note that the sine curves in Concept 1 are in phase. This is shown using a phasor diagram in Concept 2. The potential difference and current are expressed as functions of time in Equations 1 and 2.



Alternating current and a resistor
Potential difference and current in phase



Alternating current and a resistor
Potential difference, current in phasor diagram



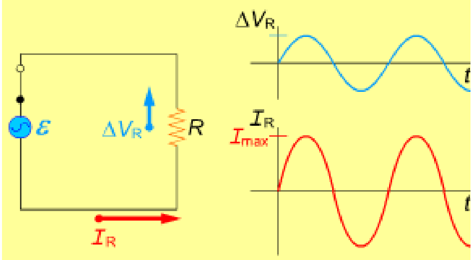
Potential difference across resistor

$$\Delta V_{\text{R}} = \Delta V_{\text{max}} \sin \omega t$$

ΔV_{R} = resistor potential difference

ΔV_{max} = maximum potential difference
$\Delta V_{max} = \mathcal{E}_{max}$
ω = angular frequency, t = time

equation 2

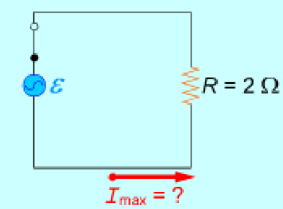


Current

$$I_R = I_{max} \sin \omega t$$

I_R = current through resistor circuit
I_{max} = maximum current
$I_{max} = \Delta V_{max} / R$
R = resistance

example 1



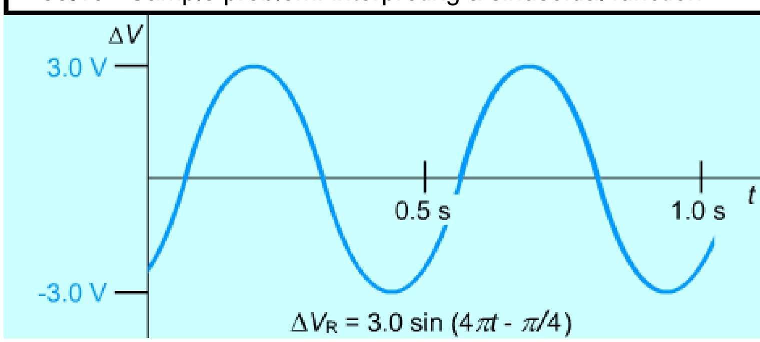
What is the maximum current in this circuit? The maximum potential difference across the resistor is 10 V.

$$I_{max} = \frac{\Delta V_{max}}{R}$$

$$I_{max} = \frac{10\ \text{V}}{2\ \Omega}$$

$$I_{max} = 5\ \text{A}$$

33.10 - Sample problem: interpreting a sinusoidal function



The potential difference across the resistor in an AC resistor circuit is $\Delta V_R = 3.0 \sin (4\pi t - \pi/4)$.

What is the maximum potential difference? What is the angular frequency? What is the initial potential difference?

Variables

potential difference across resistor	$\Delta V_R = 3.0 \sin(4\pi t - \pi/4)$ V
initial potential difference	ΔV_0
maximum potential difference	ΔV_{max}
angular frequency	ω

What is the strategy?

1. Identify the elements of the potential difference function that correspond to the desired quantities.
2. State the value that answers each question.

Step-by-step solution

The first two questions can be answered by examining the potential difference function. The maximum potential difference ΔV_{max} is the amplitude (coefficient) of the sine function: 3.0 V. The angular frequency is the coefficient of t in the argument of the sine: 4π. We find the initial potential difference by evaluating the function at $t = 0$.

Step	Reason
1. $\Delta V_R = 3.0 \sin(4\pi t - \pi/4)$ V	potential difference function
2. $\Delta V_0 = 3.0 \sin(0 - \pi/4)$ V $\Delta V_0 = -2.1$ V	evaluate at $t = 0$

33.11 - AC capacitor circuit: phase differences

The simplest AC circuit consists of just a generator and a resistor. The combination of an AC generator and a capacitor provides a slightly more complex configuration, which we now study. In this section, we focus on the relationship between the current in the circuit and the potential difference across the capacitor. As you will see, the current and potential difference functions are not in phase as they are in an AC resistor circuit.

To the right you see a circuit that consists of an AC generator and a capacitor in series. We assume that the resistance of the wire in the circuit is negligible.

We choose to begin our analysis at a moment when the current is flowing clockwise at its maximum rate and the capacitor is uncharged. The current is the same at all points in the circuit (except between the plates of the capacitor, where no current flows).

As the current flows, the two plates of the capacitor begin to charge, and the potential difference across the capacitor increases. As long as the current flows in the clockwise direction, the capacitor will continue to charge. Since we started this analysis with the alternating current at its peak, the amount of AC current is decreasing to zero over time.

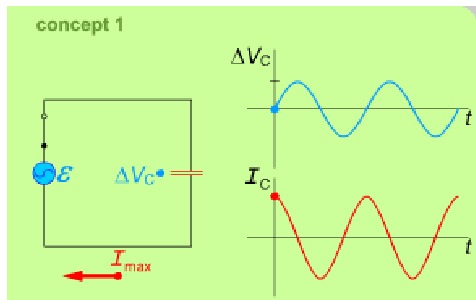
This initial part of the cycle establishes the timing relationship between the potential difference and the current. The current starts at a maximum while the potential difference across the capacitor is zero, since it is initially uncharged. When the current reaches zero, the plates of the capacitor are fully charged, and the magnitude of the potential difference across it is at a maximum. With the current flowing clockwise, the top plate of the capacitor will be positively charged.

This is unlike a circuit containing only a generator and a resistor, where the maximum potential difference across the resistor and the maximum current in the circuit occur simultaneously. Recall that the potential difference and current are in phase in that circuit.

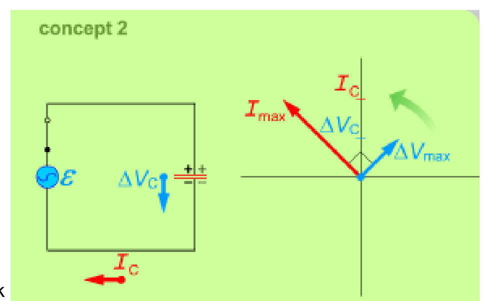
After reaching zero, the current starts to flow in the opposite direction. This causes the capacitor to discharge, and it continues to do so until no charge is left on its plates. The capacitor charged during one-quarter of a cycle of the current (when it went from a peak to zero), so it fully discharges during the next quarter cycle. In this quarter cycle, the current goes from zero to a maximum, but now flowing in the opposite direction. There is no charge on the capacitor, and no potential difference across it, at the exact moment when the current is flowing at full strength in the counterclockwise direction.

The current continues to flow in a counterclockwise fashion, and the bottom plate of the capacitor becomes positively charged. This plate continues to charge until the current again becomes zero. Now again the capacitor has its maximum charge and a maximum potential difference, but with the opposite polarity than before. This is reflected as a trough in the potential difference graph to the right.

The potential difference across the capacitor changes sinusoidally, as does the current in the circuit. You see this graphed in Concept 1. Note that the potential difference across the capacitor "follows" the current. The current reaches a peak earlier in time than the potential difference does. The expression is that "current leads potential difference." This can also be seen in the phasor diagram in Concept 2.



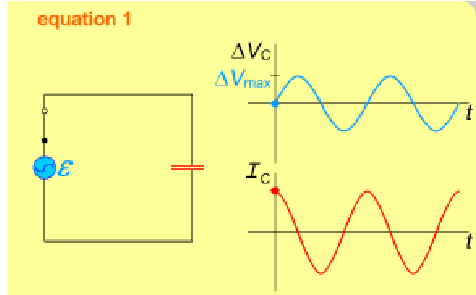
Alternating current and a capacitor
Potential difference, current not in phase
Current leads potential difference by $\pi/2$ radians



Alternating current and a capacitor
Potential difference, current in phasor diagram

The equations for potential difference and current as functions of time are shown. Note that the equation for the current has a phase constant of positive $\pi/2$ radians $(90°)$ while the phase constant for the potential difference is zero. This means there is a phase difference of $\pi/2$ radians between them, and that the current "leads" by this angle.

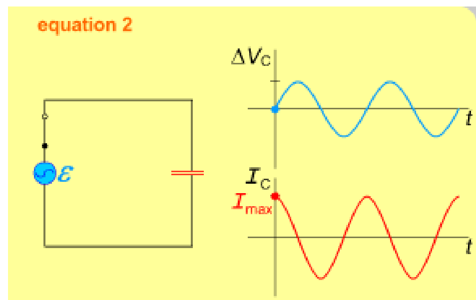
The phasors are like two horses on a carousel, always moving, and always separated by a constant distance. If you have to bet on one winning, bet on "Current" beating "Potential difference." It wins every race by a quarter lap.



Potential difference across capacitor

$$\Delta V_C = \Delta V_{max} \sin \omega t$$

ΔV_C = capacitor potential difference
ΔV_{max} = maximum potential difference
$\Delta V_{max} = \mathcal{E}_{max}$
ω = angular frequency, t = time



Current

$$I_C = I_{max} \sin (\omega t + \pi/2)$$

I_C = current through capacitor circuit
I_{max} = maximum current
$+\pi/2$ rad = phase constant

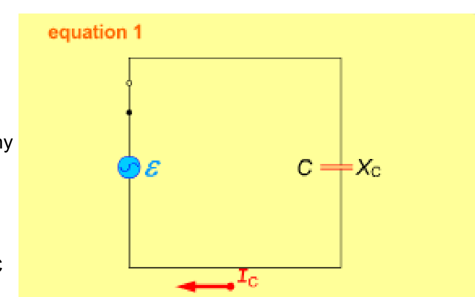
33.12 - Capacitive reactance

Just as resistance was important to analyzing an AC resistor circuit, the concept of reactance is necessary to understand the functioning of an AC capacitor circuit.

Again, Kirchhoff's loop rule can be used to analyze the circuit. Because the sum of the potential differences around the circuit must equal zero, the potential difference across the capacitor must be equal in magnitude to the emf created by the AC generator at any moment in time.

With the AC resistor circuit, we used the resistor's resistance and Ohm's law to determine the current in the circuit. Here, we must use the *capacitive reactance* of the capacitor. The capacitive reactance (X_C) equals the reciprocal of the product of the AC angular frequency and the capacitance, meaning that a greater capacitance provides less reactance. Like resistance, reactance measures the tendency of the circuit to oppose the flow of current. You see the definition in Equation 1. The units for capacitive reactance are ohms, just as they are for resistance.

The definition of capacitive reactance should appeal to your physics intuition. A greater capacitance corresponds to a greater charge on the capacitor plates for a given potential difference. (Remember: The charge equals the capacitance times the potential difference.) A greater charge, accumulating and dissipating over a time period determined by the angular frequency of the generator, results in a greater current. The capacitive reactance is inversely proportional to the capacitance: Increase the capacitance and you decrease the reactance, and more current flows.



Capacitive reactance

$$X_C = \frac{1}{\omega C}$$

X_C = capacitive reactance
ω = angular frequency of AC

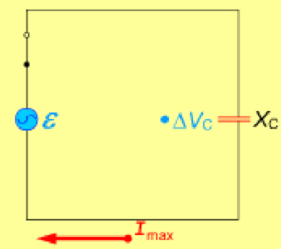
We can also consider how the angular frequency relates to capacitive reactance. If the charge on the capacitor plates accumulates and dissipates over a shorter period of time, then the current is greater. The angular frequency appears in the denominator of the defining formula. The reactance is also inversely proportional to this quantity.

In Equation 2, we show the *capacitor circuit equation*, used to relate the maximum potential difference across the capacitor to the maximum current in the circuit. The capacitor circuit equation has to be stated in terms of maximums. One **cannot** simply say, “the potential difference across the capacitor equals the product of the current and the capacitive reactance at any time.” This is because, as you have seen, the potential difference and the current are out of phase.

The maximum potential difference equals the product of the maximum current and the capacitive reactance. To put it another way, the greater the capacitance − and because of that the smaller the reactance − the larger the swings in the current will be for a given maximum potential difference. In this circuit with just an AC generator and a capacitor, the maximum potential difference equals the maximum emf of the generator.

C = capacitance
Units: ohms

equation 2

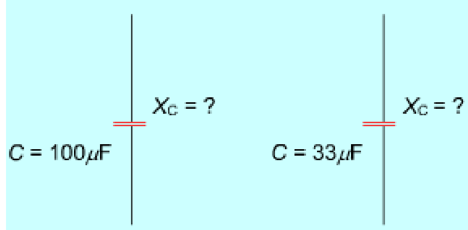


Capacitor circuit equation

$$\Delta V_{max} = I_{max} X_C$$

ΔV_{max} = maximum potential difference
$\Delta V_{max} = \mathcal{E}_{max}$
I_{max} = maximum current
$\mathcal{E}_{max}$ = maximum emf of generator

example 1



Calculate the capacitive reactance of each capacitor in a 60 Hz AC circuit.

$\omega = 2\pi f$

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C}$$

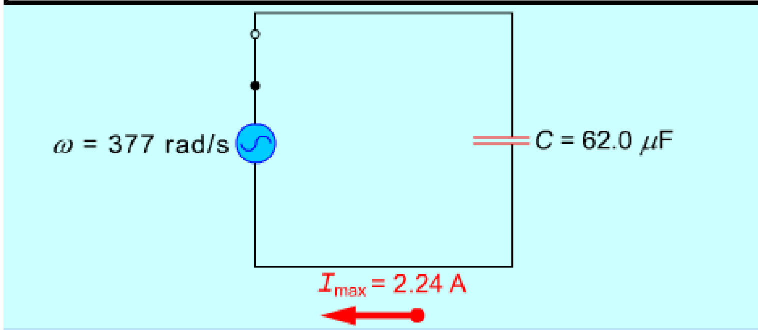
Left capacitor:

$$X_C = \frac{1}{2\pi\,(60\text{Hz})(1.0\times 10^{-4}\text{F})} = 27\,\Omega$$

Right capacitor:

$$X_C = \frac{1}{2\pi\,(60\text{Hz})(3.3\times 10^{-5}\text{F})} = 80\,\Omega$$

33.13 - Interactive checkpoint: AC capacitor circuit

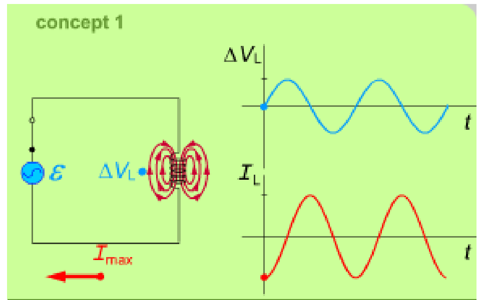


A circuit containing an AC generator and a 62.0 μF capacitor oscillates at 377 rad/s. The maximum current in the circuit is measured to be 2.24 A. What is the potential difference across the capacitor at t = 3.14 s? Assume that at t = 0 s the capacitor starts with no charge and that the current is at its positive maximum.

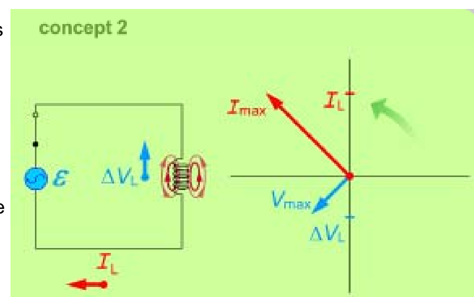
Answer:

ΔV_C = ______ V

33.14 - AC inductor circuit: phase differences



Alternating current and an inductor
Rate of change of current determines potential difference
Current lags potential difference by $\pi/2$



Alternating current and an inductor
Potential difference, current in a phasor diagram

We will now explore what happens when an inductor is placed in series with an AC generator.

To understand the nature of an inductor in an AC circuit, it is important to remember that the magnitude of the emf induced across the inductor equals the inductance times the **rate of change** of the current.

Remember that the induced emf of the inductor is oriented so it opposes the change in current. An oscilloscope would measure a potential difference equal to that emf.

To state the general precept: The inductor's potential difference always opposes change. When the current is increasing, the potential difference across the inductor will oppose the increase, and when the current is decreasing, the potential difference opposes the decrease. In short, the orientation of the potential difference across the inductor changes as the current changes.

Now let's consider this in terms of the phase relationship between the current and the potential difference across the inductor. We consider the current first, using the graph to determine the **rate of change** in the current, which equals the slope of the current graph.

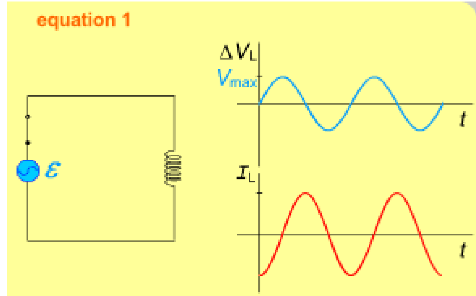
Here, we start the current at its most negative value. This is in line with the equation we use for the current. At this point, the slope of its graph equals zero, which means that its rate of change equals zero. The potential difference across the inductor is zero at this moment. The graph of the current then moves towards zero (becomes less negative): It has a positive slope. This means the potential difference is positive.

As the current changes from negative to positive, at the point where its graph crosses the horizontal axis, its slope (and rate of change) are at their maximum positive value, so the potential difference across the inductor is at its maximum positive value. Note the difference between them: The current is at zero when the potential difference is at a maximum.

The current continues to increase and it changes direction where its graph reaches a peak. At the peak, the current ceases to change for an instant so the slope again is zero; both facts explain why the potential difference is zero once again. Note that as before, one value is at a maximum when the other is zero.

Though there are many similarities to the case of a capacitor in series with an AC generator, there are also important differences. Looking at these graphs, you can see that the current "lags" the potential difference by $\pi/2$ radians $(90°)$. This is expressed in the equations: The argument of the sine function for the current is the product of the angular frequency and the time, **minus** $\pi/2$. You can also see this relationship in the phasor diagram in Concept 2. In contrast, with a capacitor, the phase constant of the current function is **plus** $\pi/2$.

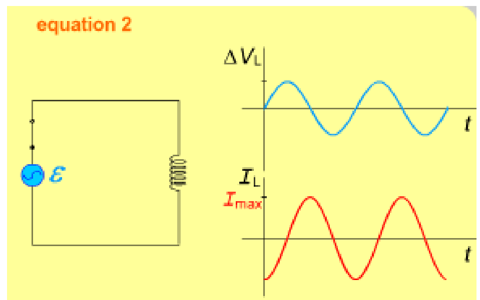
There are mnemonics for these phase relationships: ELI and ICE. ELI, in addition to being the nickname for Yalies, stands for: emf ($\mathcal{E}$) − inductor (**L**) − current (**I**). The emf in an inductor comes before the current. ICE stands for: current (**I**) − capacitor (**C**) − emf ($\mathcal{E}$). The current in a capacitor comes before the emf.



Potential difference across inductor

$$\Delta V_L = \Delta V_{max} \sin \omega t$$

ΔV_L = inductor potential difference
ΔV_{max} = maximum potential difference
$\Delta V_{max} = \mathcal{E}_{max}$
ω = angular frequency, t = time



Current

$$I_L = I_{max} \sin (\omega t - \pi/2)$$

I_L = current through inductor circuit
I_{max} = maximum current
$-\pi/2$ rad = phase constant

33.15 - Inductive reactance

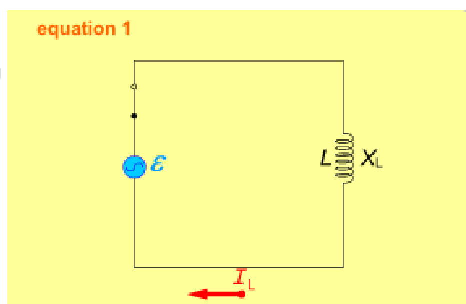
Our exploration of the role of capacitors in an AC circuit required the concept of capacitive reactance. The analogous concept for an inductor in an AC circuit is *inductive reactance*, X_L. It equals the product of the angular frequency of the alternating current and the inductance of the inductor. This is shown on the right in Equation 1. The units of inductive reactance are ohms.

In the circuit to the right, Kirchhoff's loop rule dictates that the magnitude of the potential difference across the inductor must equal the emf of the generator. The maximum potential difference equals the maximum current times the inductive reactance. This relationship is stated in Equation 2 as the *inductor circuit equation*.

As the inductor circuit equation implies, the greater the inductive reactance, the less the maximum current.

Since the inductive reactance is proportional to the inductance, this means that for a given AC generator, a greater inductance will produce a circuit with a lesser maximum current. Why? In physical terms, an inductor with greater inductance provides more opposition to changes in the flow of current, which reduces the maximum flow of the alternating current in the circuit.

Likewise, when the angular frequency of the AC is greater, the current oscillates more rapidly, and its rate of change is greater for a given amplitude. The more rapid the change, the more an inductor opposes it. In mathematical terms, the inductive



Inductive reactance

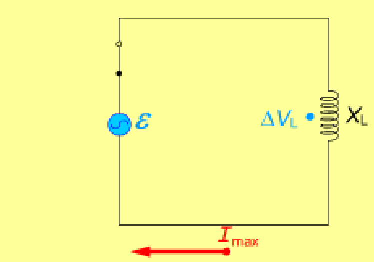
$$X_L = \omega L$$

X_L = inductive reactance
ω = angular frequency of AC

reactance is also proportional to the angular frequency of the alternating current.

L = inductance
Units: ohms

equation 2

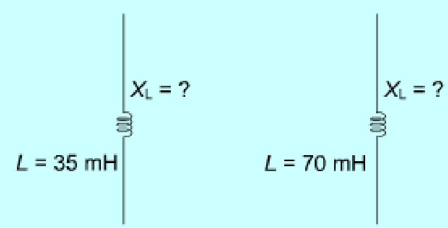


Inductor circuit equation

$$\Delta V_{max} = I_{max} X_L$$

ΔV_{max} = maximum potential difference
$\Delta V_{max} = \mathcal{E}_{max}$
I_{max} = maximum current

example 1



Calculate the inductive reactance of each inductor in a 60 Hz AC circuit.

$\omega = 2\pi f$

$X_L = \omega L = 2\pi f L$

Left inductor:

$X_L = 2\pi\,(60\text{Hz})\,(3.5 \times 10^{-2}\text{H}) = 13\,\Omega$

Right inductor:

$X_L = 2\pi\,(60\text{Hz})\,(7.0 \times 10^{-2}\text{H}) = 26\,\Omega$

33.16 - Interactive problem: angular frequency and reactance

A crucial point in this chapter is that the reactances of capacitors and inductors are partly a function of the circuit's angular frequency. This simulation is designed for you to explore this concept.

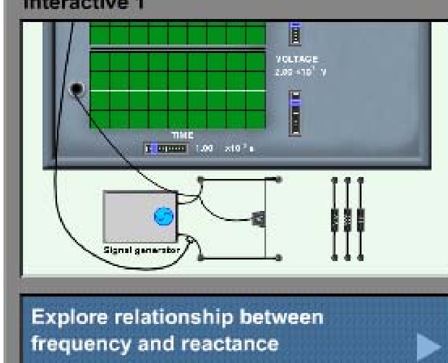
In the simulation, you have an AC signal generator in a circuit into which you can also place, one at a time, a resistor, a capacitor or an inductor.

You can vary the frequency of the driving signal using up and down arrows on the control panel. Your first task is to vary this frequency and observe the effect it has on the current in a circuit containing first a resistor, then a capacitor and finally an inductor.

You can analyze the current using an oscilloscope. Its operation was detailed in a prior section. The oscilloscope displays current in its upper display and potential difference, or voltage, at the bottom. Each display has a slider to set its scale.

If you put a resistor into the circuit, does increasing the signal frequency cause the amplitude of the alternating current to increase, decrease or stay the same? How

interactive 1



about a capacitor? An inductor?

You can answer these questions by dragging each kind of component into the circuit. Once you have explored the effects of signal frequency on that component, drag it away and put in another one.

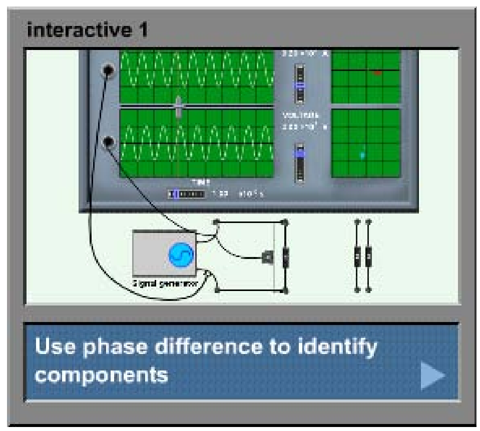
Once you have completed this experiment, your second task is to determine the capacitance of the capacitor and the inductance of the inductor.

To do so, note the frequency and then use the oscilloscope to determine the maximum current and potential difference. Equations similar in form to Ohm's law relate the reactances of capacitors and inductors to $I_{\max}$ and $\Delta V_{\max}$. Use the data from the simulation to calculate the reactance for each component at this frequency.

Two more equations relate the capacitance of a capacitor and the inductance of an inductor to reactance at a particular angular frequency. Use these equations to determine values for C, in farads (F) and L, in henrys (H). Type those values in the text entry boxes and press CHECK to see if you are correct.

You can restart this interactive at any time by pressing RESET. If you have trouble answering any of the questions, review the sections of the textbook on capacitive reactance and inductive reactance. Make sure you take into account that you are changing the frequency of the AC signal generator in the simulation, while the equations for reactance require the **angular** frequency. The angular frequency equals the frequency times 2π.

33.17 - Interactive problem: phase differences of components



The phase difference between current and potential difference is a crucial property of alternating current circuits. This interactive exercise allows you to experiment with the phase differences found in various components.

In this simulation, there is a circuit into which you can place, one at a time, any of three "mystery" electrical components. Your task is to determine which component is a resistor, which is a capacitor and which is an inductor.

The circuit is wired to an AC signal generator that is the source of the emf that causes current to flow in the circuit. Use the oscilloscope and the phase relationship between potential difference and current to determine which component is which.

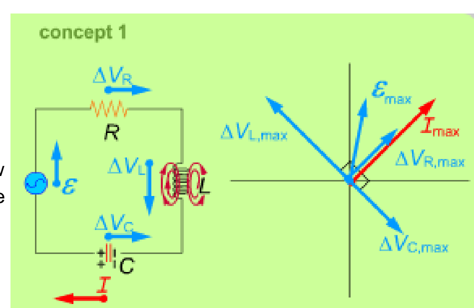
This oscilloscope features a useful tool: phasor diagrams. They make it easier to determine the phase relationship between two waves. Use the phasors to determine if the waves are in phase, and if not, whether potential difference leads current, or vice versa.

Place the components into the circuit one at a time. The oscilloscope will display the potential difference across the component and the current in the circuit. To the right of the current window is the current phasor. To the right of the voltage window is the potential difference phasor.

In the control panel, identify each component using the drop-down menus provided. When you think you have identified all three components, press CHECK to see if you are correct.

If you have trouble with this problem, review the sections of this chapter on the phase relationship between current and potential difference in an AC circuit that contains a generator and either a resistor, a capacitor, or an inductor.

33.18 - A generator and an RLC circuit: impedance



AC in a series *RLC* circuit
Resistor: current and ΔV_R in phase
Inductor: current lags ΔV_L by $\pi/2$ (90°)
Capacitor: current leads ΔV_C by $\pi/2$

The properties of a resistor, a capacitor and an inductor in an AC circuit have been analyzed separately in previous sections. By combining these three components together in series, we obtain a useful circuit that is applied in devices such as radio tuners. Such a combination of components is called a series *RLC* circuit and is the focus of this section.

To analyze the circuit, we will apply the phase relationships discussed earlier. To review them briefly: The current leads the potential difference across the capacitor and lags the potential difference across the inductor. There is no phase difference between the current and the potential difference across the resistor. It is also crucial to note that the current is the same at all points in this series circuit (excluding the gap between the plates of the capacitor, through which no current directly flows).

The basic precepts that dictate the potential differences across each component can be used to explain the timing relationships in this compound circuit. Let's again assume the current starts at a maximum. This means the potential difference across the resistor will be at a maximum, since we apply Ohm's law ($V = IR$) to determine that potential difference. The potential difference across the capacitor is zero (since it starts uncharged), and the potential difference across the inductor is also zero (since the rate of change of current is zero at a peak).

Now let's move to when the current is zero. The potential difference across the resistor is also zero, since no current is flowing through it. The magnitude of the potential difference across the capacitor is at a maximum, since it has now reached its maximum charge. The magnitude of the potential difference across the inductor is also at a maximum, since the current is changing at its maximum rate. By considering which capacitor plate is positive, and the orientation of the potential difference in the inductor that opposes the change in current, you can conclude that the potential differences have opposite orientations.

Using phasors helps to explain this kind of circuit. Multiple superimposed sinusoidal graphs are quite hard to read, as you may conclude as you view those shown on the right. The phasor diagram shows the current, the emf, and three component potential differences. The length of each phasor is the maximum value for the physical quantity that it represents.

Because the current leads the potential difference in a capacitor, the potential difference phasor for the capacitor is $\pi/2$ rad $(90°)$ behind the current phasor. Conversely, since the current lags potential difference in an inductor, the potential difference phasor for the inductor is $\pi/2$ ahead of that of the current. As the phasors rotate, these phase differences stay the same.

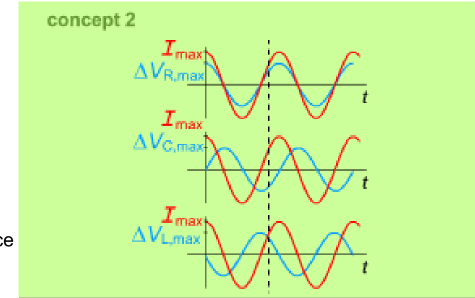
The phasors show the relationship between the potential differences across the various components at any time. For instance, consider what occurs when the phasor diagram at the right rotates so the current is at its maximum (the red phasor points straight up). This means the potential difference across the resistor is at a maximum because its phasor is vertical, as well. In contrast, the phasors representing the capacitor and the inductor will be horizontal at this time, indicating that there is no potential difference across those components.

Conversely, when the current phasor is at zero degrees, the inductor potential difference is at its positive maximum and the capacitor potential difference is at its largest negative value. The potential difference across the resistor is zero.

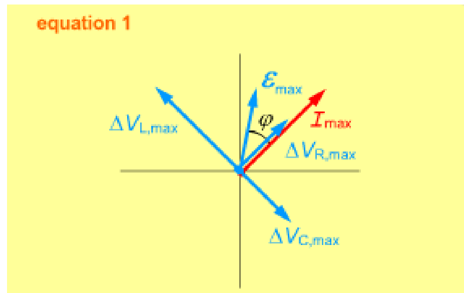
The phasor for the potential difference, or emf, provided by the AC generator is labeled $\mathcal{E}$. It is not necessarily in phase with any of the other potential difference phasors, or even with the current phasor. Kirchhoff's loop rule states that the sum of the potential differences around the circuit must be zero. This means the sum of the potential differences at any time t across the resistor, inductor and capacitor has the same magnitude as the emf supplied by the AC generator. To put it another way, the vector sum of the inductor, capacitor and resistor potential difference phasors equals the AC generator emf phasor.

The formula for calculating the phase difference between the emf phasor and the AC current in the circuit (the angle designated φ on the phasor diagram) is shown in Equation 1. It equals the arctangent of the difference between the inductive reactance X_L and the capacitive reactance X_C, divided by the resistance R. This phase difference can be positive (when $X_L > X_C$), negative (when $X_L < X_C$), or zero (when $X_L = X_C$). This means that when the reactances are equal, the emf and the current are in phase.

The equation for the *impedance* Z of the circuit is shown in Equation 2. As the formula suggests, the impedance is a combination of the individual resistive characteristics of the three components. Impedance relates potential difference to the current in the circuit, playing a role analogous to resistance in Ohm's law. The relationship of the maximum emf, maximum current and impedance is shown in Equation 2. These maximums are in phase only if the inductive and capacitive reactances are equal. As with resistance and reactance, impedance is measured in ohms.



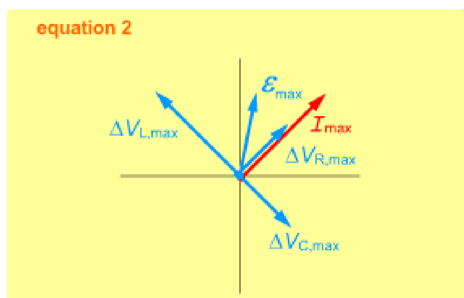
Phase differences shown in graphs



Phase difference between emf and current

$$\varphi = \arctan\frac{X_L - X_C}{R}$$

φ = emf-current phase difference
X_L = inductive reactance
X_C = capacitive reactance
R = resistance



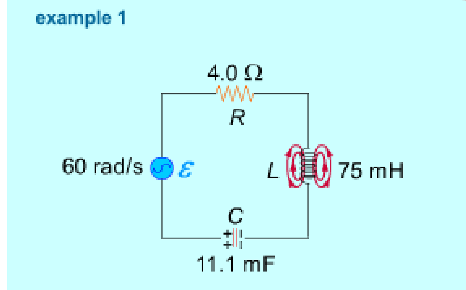
Impedance relates potential difference to current

$$\mathcal{E}_{max} = I_{max} Z$$

$\mathcal{E}_{max}$ = maximum emf in circuit
I_{max} = maximum current in circuit

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

Z = impedance of circuit
Units: ohms



What is the impedance of this circuit?

$X_L = \omega L = (60 \text{ rad/s})(0.075 \text{ H})$

$X_L = 4.5\ \Omega$

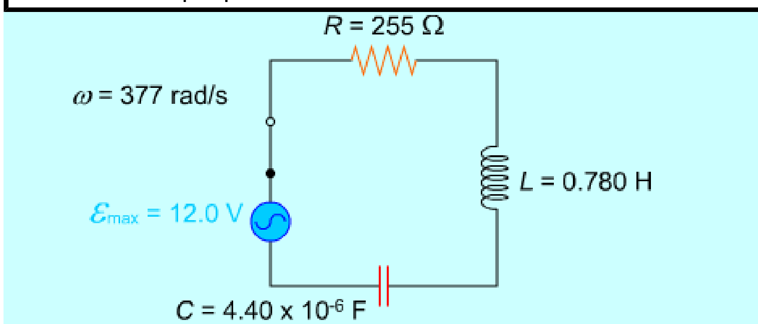
$$X_C = \frac{1}{\omega C} = \frac{1}{(60 \text{ rad/s})(0.0111 \text{ F})}$$

$X_C = 1.5\ \Omega$

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

$$Z = \sqrt{(4.0)^2 + (4.5 - 1.5)^2} = 5.0\ \Omega$$

33.19 - Sample problem: RLC maximum current



An *RLC* circuit containing a 255 Ω resistor, a 4.40x10 $^{-6}$ F capacitor, and a 0.780 H inductor is driven by a 12.0 V AC generator operating at 377 rad/s. What is the maximum current that this circuit carries?

Variables

resistance	$R = 255\ \Omega$
capacitance	$C = 4.40\times10^{-6}$ F
inductance	$L = 0.780$ H
AC emf	$\mathcal{E}_{max} = 12.0$ V
angular frequency	$\omega = 377$ rad/s
capacitive reactance	X_C
inductive reactance	X_L
impedance	Z
current amplitude	$I_{max} = ?$

What is the strategy?

1. Find the impedance of the circuit.
2. Use the equation that states the maximum current as a function of the generator emf and the impedance.
3. Substitute known quantities and evaluate.

Physics principles and equations

The equation for impedance

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

The definition of capacitive reactance

$$X_C = \frac{1}{\omega C}$$

The definition of inductive reactance

$$X_L = \omega L$$

The equation for AC circuit maximum emf and current

$$\mathcal{E}_{max} = I_{max} Z$$

Step-by-step solution

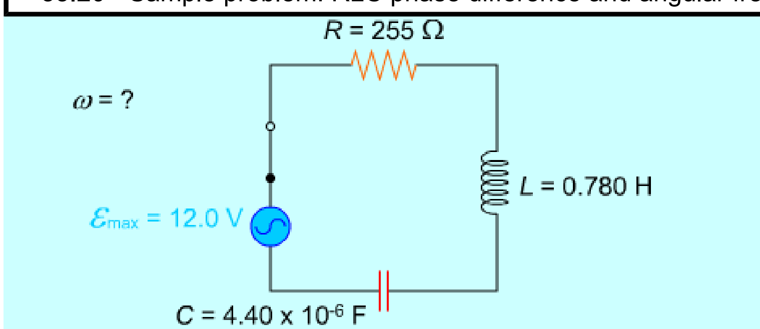
We start by finding the impedance of the circuit.

Step	Reason
1. $Z = \sqrt{R^2 + (X_L - X_C)^2}$	impedance equation
2. $X_C = \frac{1}{\omega C}$	definition of capacitive reactance
3. $X_C = \frac{1}{(377\text{ rad/s})(4.40 \times 10^{-6}\text{F})}$ $X_C = 603\ \Omega$	evaluate
4. $X_L = \omega L$	definition of inductive reactance
5. $X_L = (377\text{ rad/s})(0.780\text{ H})$ $X_L = 294\ \Omega$	evaluate
6. $Z = \sqrt{(255\ \Omega)^2 + (294\ \Omega - 603\ \Omega)^2}$ $Z = 401\ \Omega$	substitute equations 3 and 5 into equation 1

We now use the impedance to find the maximum current in the circuit.

Step	Reason
7. $I_{max} = \frac{\mathcal{E}_{max}}{Z}$	AC circuit emf related to current
8. $I_{max} = \frac{12.0\text{ V}}{(401\ \Omega)}$ $I_{max} = 2.99 \times 10^{-2}\text{A}$	substitute and evaluate

33.20 - Sample problem: RLC phase difference and angular frequency



The *RLC* circuit shown is driven by an AC generator. What frequency of oscillation will give a phase difference between the emf and current of $+\pi/4$?

Variables

resistance	$R = 255\ \Omega$
capacitance	$C = 4.40\times10^{-6}$ F
inductance	$L = 0.780$ H
AC maximum emf	$\mathcal{E}_{max} = 12.0$ V
phase difference	$\varphi = +\pi/4$
capacitive reactance	X_C
inductive reactance	X_L
angular frequency	ω

What is the strategy?

1. Start with the equation for the phase difference between the emf and current in an AC *RLC* circuit.
2. Substitute the definitions of capacitive and inductive reactance.
3. Solve for the oscillation frequency.

Physics principles and equations

The definitions of capacitive and inductive reactance are

$$X_C = \frac{1}{\omega C},\ X_L = \omega L$$

The phase difference between the emf and current in an *RLC* circuit is given by

$$\varphi = \arctan\left(\frac{X_L - X_C}{R}\right)$$

Mathematics principles

We will use the quadratic formula.

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Step-by-step solution

We begin by deriving an expression for the phase difference in terms of the capacitance and inductance, which are known, and the angular frequency, which we want to calculate.

Step	Reason
1. $\varphi = \arctan\left(\frac{X_L - X_C}{R}\right)$	phase difference
2. $X_C = \frac{1}{\omega C}$	definition of capacitive reactance
3. $X_L = \omega L$	definition of inductive reactance
4. $\varphi = \arctan\left(\frac{\omega L - \frac{1}{\omega C}}{R}\right)$	substitute equations 2 and 3 into equation 1

Next we solve the equation in step 4 for the angular frequency.

Step	Reason
5. $\tan\varphi = \dfrac{\omega L - \dfrac{1}{\omega C}}{R}$	definition of arctangent
6. $\omega L - R\tan\varphi - \dfrac{1}{\omega C} = 0$	rearrange
7. $\omega^2 L - \omega R\tan\varphi - \dfrac{1}{C} = 0$	multiply by ω
8. $\omega = \dfrac{R\tan\varphi \pm \sqrt{R^2\tan^2\varphi + \dfrac{4L}{C}}}{2L}$	quadratic formula

Now we can substitute the known values to calculate the angular frequency.

Step	Reason
9. $\omega = \dfrac{R\tan(\pi/4) \pm \sqrt{R^2\tan^2(\pi/4) + \dfrac{4L}{C}}}{2L}$ $\omega = \dfrac{R \pm \sqrt{R^2 + \dfrac{4L}{C}}}{2L}$	evaluate tangents
10. $\omega = \dfrac{(255\ \Omega) \pm \sqrt{(255\ \Omega)^2 + \dfrac{4(0.780\ \text{H})}{(4.40\times10^{-6}\ \text{F})}}}{2(0.780\ \text{H})}$	enter values
11. $\omega = \dfrac{255\ \Omega \pm 880\ \Omega}{1.56\ \text{H}}$	evaluate radical
12. $\omega = \dfrac{255\ \Omega + 880\ \Omega}{1.56\ \text{H}}$ $\omega = 728\ \text{rad/s}$	use positive solution and evaluate

33.21 - Derivation: RLC impedance and phase constant

The second equation on the right expresses the maximum current I_{max} in an AC series *RLC* circuit in terms of the maximum emf $\mathcal{E}_{max}$ and the impedance Z of the circuit. Impedance is a measure of an *RLC* circuit's tendency to limit the flow of current. It is analogous to the concepts of resistance and reactance that apply to simpler circuits.

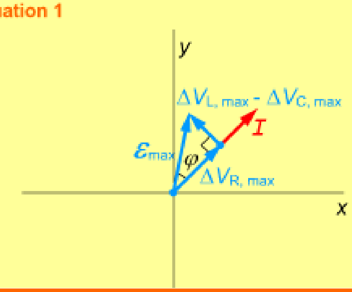
In the first equation, the impedance itself is defined in terms of the resistance, the inductive reactance, and the capacitive reactance of the components in the circuit. The third equation expresses the phase constant of the circuit. This is the phase difference between the generator emf and the AC current.

The maximum current and phase constant equations can be derived using the right triangle, known as an *impedance triangle*, which you see constructed in the phasor diagram to the right.

Kirchhoff's loop rule states that the sum of the potential differences around the circuit must be zero. This means the sum of the potential differences at any time *t* across the resistor, inductor and capacitor has the same magnitude as the emf supplied by the AC generator.

Now consider the phasor diagram. The vertical component, or height, of a phasor is the value of the physical quantity it represents at a given time. In the phasor diagram to the right, Kirchhoff's rule says that the sum of the heights of the resistor, inductor and capacitor potential difference phasors must equal the height of the emf phasor. As time passes, all the phasors rotate at the same speed, maintaining their relative sizes and positions. The heights have to add up, no matter what the angular displacement. That is, the sum of the resistor, inductor, and capacitor **vectors** has to equal the emf

equation 1



RLC impedance, maximum current and phase constant

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

$$I_{max} = \frac{\mathcal{E}_{max}}{Z}$$

$$\varphi = \arctan\frac{X_L - X_C}{R}$$

vector.

As the diagram illustrates, the inductor and capacitor phasors are π radians $(180°)$ out of phase with each other, so we can find the magnitude of their vector sum by subtracting the magnitude of one from the other.

After the inductor and capacitor phasors are combined in this way, the resistor phasor and the combined inductor-capacitor difference phasor are $\pi/2$ radians $(90°)$ out of phase with each other. For this reason, they form the legs of a right triangle (the impedance triangle). Since the emf phasor is their vector sum, it is the hypotenuse of the triangle.

Impedance and maximum current from the impedance triangle (derivation 1). We will first use the impedance triangle on the right, together with the Pythagorean theorem, to derive the maximum current equation you see as the second equation on the right. The first equation, for the impedance Z, we state as a definition.

Variables

emf of generator	$\mathcal{E}$
potential difference across resistor	ΔV_R
potential difference across inductor	ΔV_L
potential difference across capacitor	ΔV_C
current	I
resistance of resistor	R
inductive reactance of inductor	X_L
capacitive reactance of capacitor	X_C
impedance of *RLC* circuit	Z

Strategy

1. State the Pythagorean theorem for the impedance triangle. This will involve the magnitudes of the phasors forming the triangle, which in turn will equal the maximum values or amplitudes of the corresponding potential difference functions.
2. Replace each maximum potential difference by the maximum current times a variable that represents either resistance or reactance.
3. Simplify the preceding equation and define the impedance Z in a fashion that allows the equation to be written concisely.

Physics principles and equations

To make replacements in the Pythagorean theorem equation, we will use Ohm's law, and its analogs for capacitors and inductors.

$$\Delta V_R = IR$$

$$\Delta V_{C,\max} = I_{\max} X_C$$

$$\Delta V_{L,\max} = I_{\max} X_L$$

Step-by-step derivation

We use the Pythagorean theorem to relate the maximum emf to the maximum potential differences across the three components. We then replace each maximum potential difference by the maximum current times an ohmic quantity and solve for the maximum emf.

Step	Reason
1. $\mathcal{E}_{\max}^2 = (\Delta V_{R,\max})^2 + (\Delta V_{L,\max} - \Delta V_{C,\max})^2$	Pythagorean theorem
2. $\mathcal{E}_{\max}^2 = I_{\max}^2 R^2 + (I_{\max} X_L - I_{\max} X_C)^2$	substitute using Ohm-type laws
3. $\mathcal{E}_{\max} = I_{\max}\sqrt{R^2 + (X_L - X_C)^2}$	simplify

Now we define the impedance Z in a way that simplifies the preceding equation, and we solve for $I_{\max}$.

Step	Reason
4. $Z = \sqrt{R^2 + (X_L - X_C)^2}$	define impedance
5. $\mathcal{E}_{\max} = I_{\max} Z$	substitute equation 4 into equation 3
6. $I_{\max} = \dfrac{\mathcal{E}_{\max}}{Z}$	solve for $I_{\max}$

Phase constant from the impedance triangle (derivation 2). Now we will use the impedance triangle together with the definition of the tangent of an angle to derive the phase constant equation you see as the third equation on the right.

Variables. The phase constant φ represents the amount by which the generator potential difference is out of phase with the AC current in the circuit.

Strategy

1. State the definition for the tangent of the angle φ in the impedance triangle. This will involve the magnitudes of the phasors forming the triangle, which in turn equal the maximum values of the corresponding potential difference functions.
2. Replace each maximum potential difference on the right hand side of the preceding equation by the maximum current times a variable that represents either resistance or reactance, and solve the resulting equation for φ.

Mathematics principle

The tangent of an angle in a right triangle is

$$\tan\theta = \frac{\text{opposite side}}{\text{adjacent side}}$$

Step-by-step derivation

We state the definition of the tangent of φ in terms of the maximum potential differences across the three components. We then replace each maximum potential difference by the maximum current times an ohmic quantity and solve for φ.

Step	Reason
1. $\tan\varphi = \dfrac{\Delta V_{\mathrm{L,\,max}} - \Delta V_{\mathrm{C,\,max}}}{\Delta V_{\mathrm{R,\,max}}}$	definition of tangent
2. $\tan\varphi = \dfrac{I_{\max}X_{\mathrm{L}} - I_{\max}X_{\mathrm{C}}}{I_{\max}R}$	substitute using definitions of reactances, Ohm's law
3. $\tan\varphi = \dfrac{X_{\mathrm{L}} - X_{\mathrm{C}}}{R}$	simplify
4. $\varphi = \arctan\dfrac{X_{\mathrm{L}} - X_{\mathrm{C}}}{R}$	solve for φ

33.22 - Resonance frequency in a series AC RLC circuit

An *RLC* circuit has a *resonance frequency*. This frequency depends on the inductance and capacitance of the circuit. When the resonance frequency ω_n of a radio tuning circuit matches the broadcast frequency of a radio station, the circuit can "tune in" that station.

The concept of resonance frequency can be applied to any situation that exhibits harmonic motion. For instance, when you push someone on a swing, if you time your pushes correctly, you will propel the person higher and higher. The relationship of the frequency of your efforts to the natural frequency of the swing determines whether you push the person ever higher or not.

With an *RLC* circuit, it is an AC generator, a radio wave, or some other sinusoidal input that supplies the "pushes." The graph in Equation 1 shows that, as the input angular frequency ω of the AC generator (or radio wave, etc.) approaches the resonance frequency ω_n of the circuit, the amplitude of the alternating current in the circuit increases dramatically.

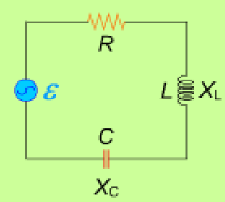
An *RLC* circuit provides the basic structure of the tuner in an old-style radio. Radio waves have different frequencies. A typical AM radio station frequency is 710 kHz (710,000 cycles per second, or 4.5 million rad/s). An *RLC* circuit can be built to have the same resonance frequency.

In radio tuners, this is usually done with a constant inductor and a variable capacitor. When you turn the tuning dial (in a radio that has one), you are changing the capacitance of a capacitor by changing the configuration of the capacitor plates. A radio can be built with the simplest of components: some wire for an antenna and the component connections, an earphone, an inductor, and a capacitor. Although the sound might be faint, this radio will work. The example problem challenges you to design a radio to receive a signal of a given frequency.

The equation for the resonance frequency in terms of the inductance and the capacitance is shown in Equation 1. When the capacitive and inductive reactances are equal, this minimizes the impedance Z of the circuit for a given resistance and maximizes the amplitude of the alternating current. Note that this resonant frequency does not depend on the resistance present in the circuit, but that R alone determines the maximum current that flows at the resonant frequency.

We derive the formula for the resonance frequency below. The variables not defined to

concept 1

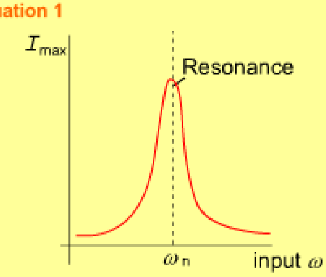


Resonance in a series *RLC* circuit

Occurs when impedance minimized

· $X_L = X_C$

equation 1



Resonance frequency of *RLC* circuit

the right are stated below.

Variables

maximum emf (amplitude) in circuit	$\mathcal{E}_{max}$
maximum current (amplitude) in circuit	I_{max}
impedance	Z
resistance	R
inductive reactance	X_L
capacitive reactance	X_C

Strategy

1. Express the maximum current in the circuit in terms of the impedance and the maximum emf of the circuit. Observe that I_{max} is maximized when the impedance is minimized.
2. Write the formula for the impedance, and observe what values of the variables in the formula will minimize the impedance. Then write an equation that has to be satisfied to minimize the impedance.
3. Solve the equation for the angular frequency that corresponds to the minimum impedance. This is the resonance frequency.

Equations

In an *RLC* circuit

$$I_{max} = \frac{\mathcal{E}_{max}}{Z}$$

The definition of impedance

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

The equations for the capacitive and inductive reactances

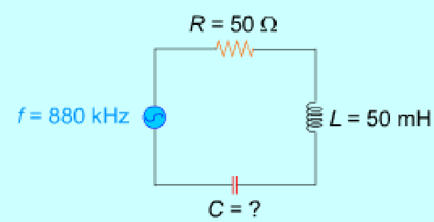
$$X_C = \frac{1}{\omega C}, \; X_L = \omega L$$

Step-by-step derivation

$$\boxed{\omega_n = \frac{1}{\sqrt{LC}}}$$

ω_n = resonance frequency
L = impedance
C = capacitance
Units: rad/s

example 1



What capacitance is required in the above circuit to tune in a station broadcasting at 880 kHz on the AM radio dial?

$$\omega_n = \frac{1}{\sqrt{LC}}$$

$$C = \frac{1}{L\omega_n^2}$$

$$C = \frac{1}{(0.050\text{ H})(2\pi \cdot 880{,}000\text{ Hz})^2}$$

$$C = 0.65\text{ pF}$$

We begin by stating the relationship between maximum current, maximum emf, and impedance. I_{max} is maximized when the impedance Z is minimized. In the subsequent steps we explore how to minimize Z.

Step	Reason
1. $I_{max} = \frac{\mathcal{E}_{max}}{Z}$	equation stated above
2. $Z = \sqrt{R^2 + (X_L - X_C)^2}$	definition of impedance
3. $X_L = X_C$ to minimize	inspection
4. $\omega_n L = \frac{1}{\omega_n C}$	substitute definitions of inductive, capacitive reactances
5. $\omega_n = \frac{1}{\sqrt{LC}}$	solve for ω_n

33.23 - Rms power in AC resistor circuits

Calculating a useful measure for power in an AC circuit is more complicated than in a DC circuit. Factors like the emf and current vary over time.

If you want to calculate the instantaneous power, you can do so by multiplying the emf and current at any instant in time. However, the *average power* of the circuit is of more interest. For instance, with a device as simple as a light bulb, it is more interesting to know the average power of the device rather than how its power varies every thousandth of a second. Knowing the average power proves useful if you want to calculate the energy consumed by an appliance or other device, since the energy will equal the product of the average power and the elapsed time.

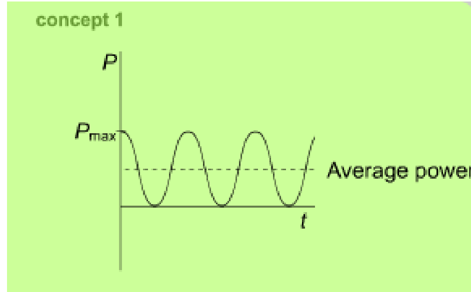
Here, we discuss how to calculate the average power in a circuit that contains solely a resistor (a light bulb) and an AC generator. To do so, we must calculate the average of the product of the emf and the current. This is different than calculating the product of their averages! Why? Because the average current and average emf equal zero: They fluctuate between peak and trough as sinusoidal waves centered on zero. However, their average product is **not** zero, since the current peak occurs at the same time as the emf peak in this circuit, and trough occurs at the same time as trough. Multiplying two negative values (at the troughs) yields a positive value for the power.

To avoid the “zero-average” difficulty mentioned above, physicists and engineers can use a specific kind of average, called the root mean square (rms) average, to calculate more representative values for the average current and emf. To calculate an rms value, you square the sinusoidal function representing the physical quantity, find the average value of the (always positive) squared function, and then take the square root of the result to compensate for squaring in the first step.

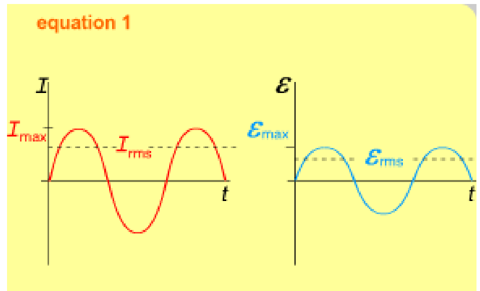
In Equation 1 you see formulas for the rms averages of the sinusoidal functions that represent current and emf. They are equal to the maximum values (amplitudes) of the functions, divided by the square root of two. These equations can be derived mathematically.

Rms values are typically used in analyzing AC circuits. For instance, ammeters and voltmeters display the rms values for current and potential difference. Electrical outlets in the United States are rated at 120 V, their rms potential difference. The peak instantaneous potential difference across the contacts is approximately 170 V.

When the rms values for the current and emf are multiplied together to determine the average power of an AC circuit, the result is the product of the maximum current and the maximum emf, divided by two. There is an important assumption here: The current and the potential difference are in phase. This is true for the case of an AC generator and a resistor (like our example of the light bulb).



Average power
Instantaneous power fluctuates in AC circuits
Average power is useful measure

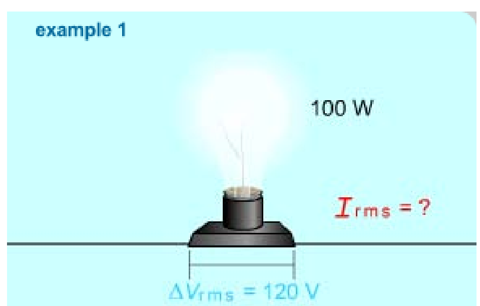


Average power in AC generator-resistor circuit

$$I_{rms} = \frac{I_{max}}{\sqrt{2}} \qquad \mathcal{E}_{rms} = \frac{\mathcal{E}_{max}}{\sqrt{2}}$$

$$P_{avg} = I_{rms}\mathcal{E}_{rms} = \frac{I_{max}\mathcal{E}_{max}}{2}$$

I_{rms} = root mean square (rms) current
I_{max} = maximum current
$\mathcal{E}_{rms}$ = rms emf
$\mathcal{E}_{max}$ = maximum emf
P_{avg} = average power



What is the rms current flowing through a 100 watt light bulb?

$P_{avg} = I_{rms}\Delta V_{rms} = I_{rms}\mathcal{E}_{rms}$
$I_{rms} = P_{avg} / \mathcal{E}_{rms}$
I_{rms} = 100 W / 120 V
I_{rms} = 0.83 A

33.24 - Power in AC RLC circuits

In this section, we delve into the topic of power in an *RLC* circuit in more detail. We will start with an energy issue: Which components dissipate the energy in an *RLC* circuit? Let's consider the role of the inductor and the capacitor in terms of energy. An important point is that in an *RLC* circuit, neither the inductor nor the capacitor is a net "consumer" of energy.

Why are these components **not** net consumers of energy? Consider again the energy of an oscillating circuit that contains only an inductor and a capacitor, with no external source of emf. The sum of the energies of these two components remains constant; it flows back and forth between them. Since the energy remains constant, these components do not dissipate energy like a resistor does.

When an AC generator is added to the circuit, you can think of the capacitor and the inductor as storing energy from or returning energy to the circuit at various times, with their net effect totaling zero. For instance, when the capacitor is being charged, energy is added to it, but when it discharges, it returns energy back to the circuit. So, too, with the inductor. It opposes the increases in current, but it also opposes the decreases. The net energy consumption of each component during a complete cycle is zero.

In contrast, a resistor does dissipate energy. It radiates heat, draining energy from the circuit. In order to analyze the energy dissipation of a circuit during a complete cycle, the inductor and capacitor can be ignored, but the resistor must be considered.

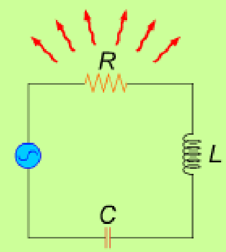
With these principles established, we move to explaining the equations for power in the *RLC* circuit shown on the right. We must factor in the timing relationships present in an *RLC* circuit. The first equation states that the **instantaneous** power equals the maximum current squared times the resistance times a sine squared function. The leading factors may look familiar: $P = I^2R$ is one equation for power in a DC circuit. The sine squared factor includes φ, the phase difference between the current in the circuit and the emf supplied by the AC generator.

In Equation 2, we show two formulas for the **average** power, each one based on the same two attributes of the circuit: the current passing through it and its resistance. These average power equations can be derived from the equation we just discussed, using some trigonometric know-how. In the first of them, the average power equals the square of the rms current times the resistance. This is one instance of how rms current helps to effectively describe a circuit.

These equations differ from the average power equation for an AC resistor-only circuit, in which the emf and the current are in phase.

The expression for average power in Equation 3 uses the angle φ. The cosine of this phase difference is called the *power factor*. When the cosine is close to zero, which indicates that φ is close to $\pi/2$ and the current is far out of phase with the emf, very little current flows through the circuit and it dissipates little power. When the power factor is one (this will occur when the capacitive and inductive reactances are equal) the circuit is operating at its resonance frequency. This equation shows the importance of capacitance and inductance in determining the power of an *RLC* circuit.

concept 1

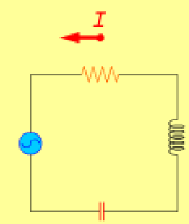


Energy/power in AC *RLC* circuits

Energy is dissipated by resistor

· Capacitor, inductor dissipate none

equation 1

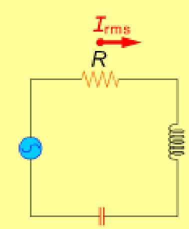


Instantaneous power

$$P = I_{max}^2 R \sin^2(\omega t - \varphi)$$

P = instantaneous power
I_{max} = maximum current (amplitude)
R = resistance
ω = angular frequency, t = time
φ = emf-current phase difference

equation 2



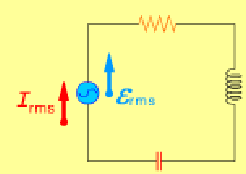
Average power as function of current, resistance

$$P_{avg} = I_{rms}^2 R$$

$$P_{avg} = \frac{I_{max}^2 R}{2}$$

P_{avg} = average power
I_{max} = maximum current
I_{rms} = root mean square current
R = resistance

equation 3

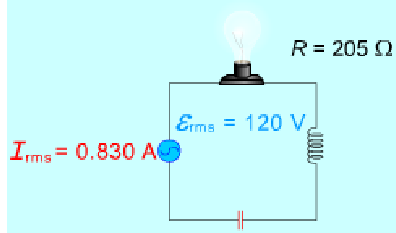


Average power as function of emf

$$P_{avg} = I_{rms}\mathcal{E}_{rms} \cos \varphi$$

$\mathcal{E}_{rms}$ = rms emf

example 1



What is the average power consumed by the light bulb in this circuit?

$P_{avg} = I_{rms}^2 R$

P_{avg} = (0.830 A)^2(205 Ω)

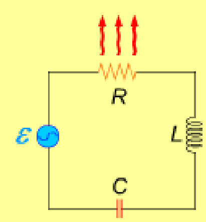
P_{avg} = 141 W

33.25 - Derivation: power equations in AC circuits

To the right you see the equations that figure in a discussion of instantaneous and average power dissipation in an AC *RLC* circuit. We will derive all the equations below. The emf supplied by the generator is $\mathcal{E} = \mathcal{E}_{max} \sin \omega t$, and the phase difference between the emf and the current in the circuit is φ.

Instantaneous power dissipation (derivation 1). In the first derivation we obtain the first equation on the right, expressing the instantaneous rate of power dissipation in the circuit as a function of time.

equation 1



Power equations for AC *RLC* circuit

$$P = I_{max}^2 R \sin^2(\omega t - \varphi)$$

$$P_{avg} = \frac{I_{max}^2 R}{2}$$

$$P_{avg} = I_{rms}^2 R$$

$$P_{avg} = I_{rms}\mathcal{E}_{rms} \cos \varphi$$

Variables

	variable	amplitude	rms average
current	I	I_{max}	I_{rms}
emf	$\mathcal{E}$	$\mathcal{E}_{max}$	$\mathcal{E}_{rms}$

instantaneous power dissipation	P
average power dissipation	P_{avg}
resistance of resistor	R
impedance of circuit	Z
angular frequency of AC	ω
time	t
emf-current phase difference	φ
resistor max potential difference	$\Delta V_{R,max}$

Strategy

1. State an equation for power dissipation, depending on the current through a circuit and the resistance of the circuit.
2. Replace the current with the function that describes it as a sinusoidal function of time.

Physics principles and equations

An equation for the power dissipation of a circuit

$$P = I^2 R$$

Since the current differs from the emf by a phase constant φ, we can write

$$I = I_{max}\sin(\omega t - \varphi)$$

Step-by-step derivation

We state the equation for power, and substitute a sinusoidal function for the current.

Step	Reason
1. $P = I^2 R$	power equation
2. $I = I_{max} \sin(\omega t - \varphi)$	current as function of time
3. $P = I_{max}^2 R \sin^2(\omega t - \varphi)$	substitute equation 2 into equation 1

Average power based on current and resistance (derivation 2). Now we proceed to derive the first two average power equations on the right, expressing P_{avg} in terms of current and resistance.

Strategy

1. Start with the instantaneous power equation. Restate the equation so it is a single expression squared.
2. Use the formula stated below for the mean square of a sine function to remove the sine squared factor. This is the first average power equation on the right.
3. Use the relationship of rms to maximum current to express the equation just derived in terms of maximum current. This is the second average power equation on the right.

Physics principles and equations

For the function $y = y_{max} \sin\theta$, the mean square of the function is

$$(y^2)_{avg} = \frac{y_{max}^2}{2}$$

The relationship between rms current and current amplitude

$$I_{rms} = \frac{I_{max}}{\sqrt{2}}$$

Step-by-step derivation

In the following steps, we restate the instantaneous power equation, and apply the formulas for the mean square of a sine function and the rms current.

Step	Reason
1. $P = ([I_{max}\sqrt{R}\,]\sin(\omega t - \varphi))^2$	instantaneous power equation rewritten
2. $P_{avg} = \frac{[I_{max}\sqrt{R}\,]^2}{2}$	mean square based on amplitude
3. $P_{avg} = \frac{I_{max}^2 R}{2}$	simplify
4. $\frac{I_{max}^2}{2} = I_{rms}^2$	rms current and current amplitude
5. $P_{avg} = I_{rms}^2 R$	substitute equation 4 into equation 3

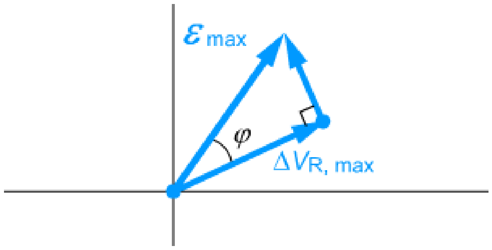
Average power as a function of emf (derivation 3). We derive the version of the average power equation that relates power dissipation to the emf of the generator, and which contains the power factor, $\cos\varphi$.

Strategy

1. It is easiest to derive this average power formula "backwards." That is, start with the final formula and expand several of the factors that appear in it. Then, simplifying at the end, you will see that it equals the average power.
2. Expand the "power factor" $\cos\varphi$ in the formula by referring to the impedance triangle. This is the key step in the derivation.
3. In the resulting trigonometric ratio, use Ohm's law and the analogous maximum current equation to replace the numerator and denominator.
4. Finally, expand the $\mathcal{E}_{rms}$ factor in the formula. Simplify the result to see that the whole formula does indeed represent the average power dissipation of the circuit.

Physics principles and equations

The potential differences across the components in the circuit can be represented graphically by the impedance triangle.



We will use Ohm's law for a resistor and the equation that relates maximum emf, maximum current and impedance.

$$\Delta V = IR, \quad \mathcal{E}_{max} = I_{max}Z$$

We can rewrite the maximum current equation as

$$\mathcal{E}_{rms} = I_{rms}Z$$

because $\mathcal{E}_{max} = \mathcal{E}_{rms}\sqrt{2}$ and $I_{max} = I_{rms}\sqrt{2}$

Step-by-step derivation

In the following steps we expand the final average power formula by using the definition of the cosine and applying several versions of Ohm's law. The result is that the formula is shown to be equal to the average power.

Step	Reason
1. $\text{formula} = I_{rms}\mathcal{E}_{rms}\cos\varphi$	formula for average power
2. $\text{formula} = I_{rms}\mathcal{E}_{rms}\dfrac{\Delta V_{R,\,max}}{\mathcal{E}_{max}}$	define cosine from impedance triangle
3. $\text{formula} = I_{rms}\mathcal{E}_{rms}\dfrac{I_{max}R}{I_{max}Z}$	Ohm's law; impedance equation stated above
4. $\text{formula} = I_{rms}(I_{rms}Z)\dfrac{I_{max}R}{I_{max}Z}$	second impedance equation stated above
5. $\text{formula} = I_{rms}^2 R = P_{avg}$	simplify

33.26 - Interactive problem: RLC radio tuner

You are trying to tune in your favorite radio show, "Physics News Tonight." You know it transmits at a broadcast frequency of 6.00×10^5 Hz (600 kHz).

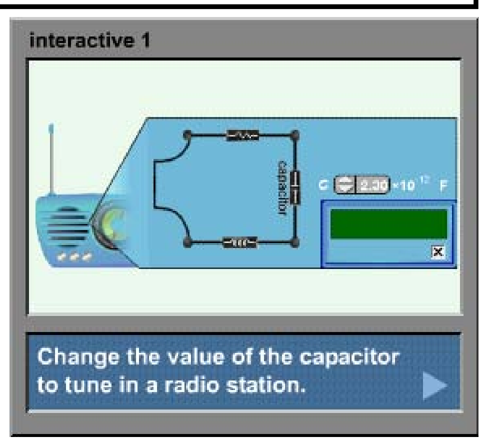
You have an *RLC* circuit like the one shown to the right. It contains an inductor with a fixed inductance of 0.00800 H. It also contains a variable capacitor. Your task is to set the capacitance of the capacitor so that the circuit has a resonant frequency of 6.00×10^5 Hz.

Calculate your answer to three significant digits. Specify the capacitance using the up and down arrows next to the capacitor in the circuit.

If you have selected the correct capacitance, you will hear a brief broadcast. If you did not, you will hear only static.

If you fail, redo your calculations and try again. Keep trying, or you will miss the broadcast and be out of touch on our version of the latest developments in the world of physics!

If you have trouble finding the right capacitance, review the section in this textbook on resonance frequency in AC *RLC* circuits.



33.27 - Gotchas

What lags and what leads? In capacitors, the current leads potential difference. In inductors, current lags. In resistors, there is no phase difference.

In the circuits discussed in this chapter, the potential differences across components at a moment in time can differ. Yes. These are series circuits; their potential differences must **sum** to zero.

In the circuits discussed in this chapter, the current flowing through the components at a moment in time can differ. No, these are series circuits, so the current is everywhere the same, except that it does not flow between the plates of the capacitors.

An AC generator is combined with a resistor, inductor and capacitor to form a series circuit. The generator determines the frequency. Yes. The other components of the circuit determine the maximum current given the maximum emf of the generator. The AC generator alone determines the frequency.

33.28 - Summary

In a series inductor-capacitor (*LC*) circuit, energy in the circuit is alternately stored in the capacitor's electric field and the inductor's magnetic field. Sinusoidal functions are used to describe the current in the circuit and the charge on the capacitor.

Adding a resistor in series to an *LC* circuit creates a series *RLC* circuit.

An AC generator creates a sinusoidal emf that drives an alternating current. The current can be described using a sine function; it flows back and forth in a circuit.

Phasors represent sinusoidal functions as vectors that rotate about the origin. The length of a phasor is equal to the amplitude of the function, which is a constant. The angle the phasor makes with the x axis is the argument of the function. The y value of the phasor is the value of the function at a particular time. The angle between two phasors equals their phase difference.

When an AC generator is connected to a resistor, the potential difference and current are in phase.

When an AC generator is connected to a capacitor, the potential difference and current are out of phase. The current **leads** the potential difference by $\pi/2$ radians. The capacitive reactance relates the maximum potential difference in the circuit to the maximum current. It equals the reciprocal of the product of AC angular frequency and the capacitance. Capacitive reactance is measured in ohms.

The potential difference and current are also out of phase when an AC generator is connected to an inductor. The current **lags** (follows) the potential difference by $\pi/2$ radians. Inductive reactance relates the maximum potential difference to the maximum current. It equals the AC angular frequency times the inductance, and is measured in ohms.

The emf phasor of the AC generator must equal the vector sum of the potential difference phasors of the resistor, inductor and capacitor in a series AC *RLC* circuit. The phase difference between the emf and current phasors is a function of the resistance, inductive reactance, and capacitive reactance of the components.

The total impedance of a series AC *RLC* circuit is a function of the resistive and reactive characteristics of the components. The impedance relates the maximum AC generator emf to the maximum current in the circuit.

The resonance frequency ω_n of an *RLC* circuit occurs when impedance is minimized. It equals the reciprocal of the square root of the product of the inductance and the capacitance.

For many applications it is useful to consider the rms average power. This average equals one-half the product of the maximum current and maximum emf for an AC resistor circuit.

The equations for power in an AC *RLC* circuit are shown to the right.

Equations

LC circuit

$$q = q_{max} \sin(\omega t + \pi/2)$$

$$I = -I_{max} \sin \omega t$$

$$I_{max} = \omega q_{max}$$

$$\omega = \frac{1}{\sqrt{LC}}$$

AC generator

$$\mathcal{E} = \mathcal{E}_{max} \sin \omega t$$

AC resistor circuit

$$\Delta V_R = \Delta V_{max} \sin \omega t$$

$$I_R = I_{max} \sin \omega t$$

$$I_{rms} = \frac{I_{max}}{\sqrt{2}} \qquad \mathcal{E}_{rms} = \frac{\mathcal{E}_{max}}{\sqrt{2}}$$

$$P_{avg} = I_{rms}\mathcal{E}_{rms} = \frac{I_{max}\mathcal{E}_{max}}{2}$$

AC capacitor circuit

$$\Delta V_C = \Delta V_{max} \sin \omega t$$

$$I_C = I_{max} \sin(\omega t + \pi/2)$$

$$X_C = \frac{1}{\omega C}$$

$$\Delta V_{max} = I_{max} X_C$$

AC inductor circuit

$$\Delta V_L = \Delta V_{max} \sin \omega t$$

$$I_L = I_{max} \sin(\omega t - \pi/2)$$

$$X_L = \omega L$$

$$\Delta V_{max} = I_{max} X_L$$

AC *RLC* circuit

$$\varphi = \arctan \frac{X_L - X_C}{R}$$

$$\mathcal{E}_{max} = I_{max} Z$$

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

$$\omega_n = \frac{1}{\sqrt{LC}}$$

$$P = I_{max}^2 R \sin^2(\omega t - \varphi)$$

$$P_{avg} = I_{rms}^2 R = \frac{I_{max}^2 R}{2}$$

$$P_{avg} = I_{rms}\mathcal{E}_{rms} \cos \varphi$$

Chapter 33 Problems

Conceptual Problems

C.1 In a low frequency AC circuit, a capacitor acts like an open circuit (extremely high impedance). However, in a high frequency circuit a capacitor behaves like a short circuit (negligible impedance). Explain this behavior.

C.2 In a low frequency AC circuit, an inductor acts like a short circuit (negligible impedance). In a high frequency circuit an inductor behaves like an open circuit (extremely high impedance). Explain this behavior.

C.3 Do reactances in a series AC circuit add like capacitors in series or do they add like resistors in series? Explain your choice.

○ Capacitors ○ Resistors

C.4 In a series *RLC* circuit, can the average power consumed by the resistor ever be negative? Explain.

○ Yes ○ No

Section Problems

Section 2 - LC oscillations quantified

2.1 A 2.5×10^{-5} F capacitor and a 2.8×10^{-2} H inductor are wired in series. The capacitor has a charge across its plates when the components are connected in a simple loop circuit at time t = 0 s. The maximum current in the resulting circuit is 7.5 A. What is the current at time t = 1.6 s?

__________ A

2.2 A simple loop circuit contains an inductor with inductance 2.45×10^{-2} H and a capacitor. (a) If the current in the circuit oscillates with a frequency of 250 Hz, what is the capacitance of the capacitor? (Note: The frequency of the circuit is stated, not its angular frequency.) (b) If the maximum charge on the capacitor is 3.28×10^{-3} C, what is the maximum current through the circuit?

(a) __________ F
(b) __________ A

2.3 A 7.80×10^{-2} H inductor and 3.30×10^{-5} F capacitor are connected in series in a simple loop. (a) What is the angular frequency of the circuit oscillations? (b) What is the frequency?

(a) __________ rad/s
(b) __________ Hz

2.4 A 4.7×10^{-6} F capacitor has a charge of 5.6×10^{-6} C. It is connected in series with a 2.9×10^{-2} H inductor and at time t = 0 s, a circuit is closed containing just these two components. What is the charge across the capacitor at time t = 4.2 ms?

__________ C

2.5 A simple loop circuit has an inductor with inductance 4.6×10^{-3} H and a capacitor with capacitance 2.8×10^{-6} F in series. At a certain time, the current is at its maximum, 5.1 A. (a) What is the charge on the capacitor at this time? (b) What is the maximum charge on the capacitor?

(a) __________ C
(b) __________ C

2.6 A 0.0200 H inductor is in series with a 3.00×10^{-11} F capacitor and a 7.00×10^{-11} F capacitor in a simple loop circuit. What is the angular frequency of this *LC* circuit?

__________ rad/s

Section 5 - An AC generator, its emf, and AC current

5.1 An AC generator generates a maximum emf of 35 V and runs at a frequency of 12 Hz. At a time t, the emf is at a maximum. How many seconds after t does it take for the emf to fall to zero?

__________ s

5.2 An AC generator is connected in series with a resistor in a simple loop circuit. At a particular time t, the current through the resistor has its maximum value of 12.3 A, and then begins to fall. Just 0.00573 s later, the current through the resistance has fallen to 10.7 A. What is the angular frequency of the generator?

__________ rad/s

Section 6 - Interactive problem: alternating current

6.1 Use the oscilloscope in the simulation in the interactive problem in this section to determine (a) I_{max}, (b) ΔV_{max} and (c) the frequency *f* of the **first** signal.

(a) _________ A
(b) _________ V
(c) _________ Hz

6.2 Use the oscilloscope in the simulation in the interactive problem in this section to determine (a) I_{max}, (b) ΔV_{max} and (c) the frequency *f* of the **second** signal.

(a) _________ A
(b) _________ V
(c) _________ Hz

Section 9 - Alternating current and a resistor

9.1 The generator in an AC resistor circuit creates a maximum potential difference across the 175 Ω resistor of 243 V. What is the amplitude of the alternating current if the frequency of the generator is (a) 455 Hz? (b) 875 Hz?

(a) _________ A
(b) _________ A

9.2 In an AC resistor circuit, the maximum potential difference across the resistor is 375 V, and the maximum current is 16.8 A. What is the resistance of the resistor?

_________ Ω

Section 10 - Sample problem: interpreting a sinusoidal function

10.1 The current through the resistor in an AC resistor circuit is defined by an equation of the form $I = I_{max}\sin(\omega t + \varphi)$. What value for φ will result in the current having a positive maximum at time $t = 0$ s?

○ 0 ○ $\pi/2$ ○ π ○ 2π

10.2 An AC resistor circuit has maximum current of 22 A. The current is −22 A at time $t = 0$ s, and it reaches zero for the first time at $t = 0.50$ s. Which equation describes the current in the circuit?

○ $I_R = (22\text{ A})\sin(\pi t)$
○ $I_R = (22\text{ A})\sin(t - \pi/2)$
○ $I_R = (22\text{ A})\sin(\pi t - \pi/2)$
○ $I_R = (22\text{ A})\sin(\pi t - \pi)$

Section 11 - AC capacitor circuit: phase differences

11.1 At time $t = 0$ s, the potential difference across the capacitor in an AC capacitor circuit is 0 V, and it is increasing. The angular frequency of the AC generator is 18.5 rad/s. What is the first time when the current in the circuit reaches 0 A?

_________ s

11.2 The AC generator in an AC capacitor circuit creates a maximum potential difference across the capacitor of 175 V and a maximum current of 22.8 A. The frequency of the generator is 977 Hz. If the potential difference across the capacitor at a time *t* is 107 V, what is the current at the same time?

_________ A

Section 12 - Capacitive reactance

12.1 The generator in an AC capacitor circuit creates a maximum potential difference across the 3.37×10^{-6} F capacitor of 243 V. What is the amplitude of the alternating current if the frequency of the generator is (a) 455 Hz? (b) 875 Hz?

(a) _________ A
(b) _________ A

12.2 What is the capacitive reactance of two square pieces of sheet metal 2.00 meters on a side, separated by 20.0 cm, which operate at an angular frequency of 3.00e+6 rad/s?

_________ Ω

Section 14 - AC inductor circuit: phase differences

14.1 At time t = 0 s, the potential difference across the inductor in an AC inductor circuit is 0 V, and it is increasing. The angular frequency of the AC generator is 38.0 rad/s. What is the first time when the current in the circuit reaches 0 A?

__________ s

14.2 The maximum potential difference across the inductor in an AC inductor circuit is 213 V and the maximum current is 18.3 A. The frequency of the AC generator is 552 Hz. If the potential difference across the inductor at a time t is 115 V, what is the current at the same time?

__________ A

Section 15 - Inductive reactance

15.1 The generator in an AC inductor circuit creates a maximum potential difference across the 2.74×10^{-3} H inductor of 243 V. What is the amplitude of the alternating current if the frequency of the generator is (a) 455 Hz? (b) 875 Hz?

(a) __________ A
(b) __________ A

15.2 A solenoid is 10.0 cm long, has a radius of 0.500 cm and has 8000 equally spaced loops. What angular frequency needs to drive this inductor so that it has a reactance of 3.00e+3 ohms?

__________ rad/s

Section 16 - Interactive problem: angular frequency and reactance

16.1 Use the simulation in the interactive problem in this section to determine (a) the capacitance of the capacitor and (b) the inductance of the inductor. Test your answer using the simulation.

(a) __________ F
(b) __________ H

Section 17 - Interactive problem: phase differences of components

17.1 Use the simulation in the interactive problem in this section determine the identity of the three "mystery" electrical components. What is the identity of component number (a) 1, (b) 2, (c) 3?

(a)
i. Inductor
ii. Capacitor
iii. Resistor

(b)
i. Inductor
ii. Capacitor
iii. Resistor

(c)
i. Inductor
ii. Capacitor
iii. Resistor

Section 18 - A generator and an RLC circuit: impedance

18.1 What is the relative phase between the voltage and the current of an AC circuit that has a 7000 ohm resistor in series with a 0.860 μF capacitor and a 0.300 H inductor? The circuit is being driven by an alternating voltage source, at an angular frequency of 600 rad/s.

__________ rad

18.2 What angular frequency is required to give a relative phase of 1.10 radians between the voltage and the current in a circuit with a 810 ohm resistor in series with a 9.00×10^{-11} F capacitor and a 1.50 H inductor?

__________ rad/s

18.3 Many radios are able to pick up different radio stations by using a capacitor whose capacitance can be changed by rotating a knob. The angular frequency of the circuit then matches the broadcast frequency. All AM radio stations broadcast between 5.30×10^{5} Hz and 1.71×10^{6} Hz. If the receiver uses a 5.00×10^{-3} H inductor, what is the range of capacitance needed to receive all possible AM radio stations?

Capacitance corresponding to higher frequency: __________ F
Capacitance corresponding to lower frequency: __________ F

Section 21 - Derivation: RLC impedance and phase constant

21.1 What is the impedance of a circuit that has a 10500 ohm resistor in series with a 3.00×10^{-8} F capacitor and a 0.500 H inductor, operating at a frequency of 500 rad/s?

__________ Ω

21.2 In a driven RLC circuit, how does the maximum current depend on the frequency in (a) the low frequency ($\omega << (LC)^{-1/2}$) regime? (b) the high frequency ($\omega >> (LC)^{-1/2}$) regime? (Hint: In these regimes, the resistance makes a negligible contribution to the total impedance.)

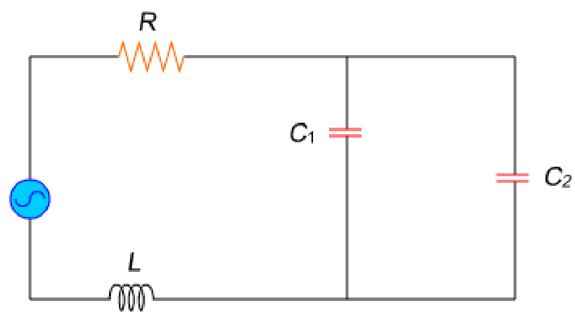
(a) The maximum current is
- i. linearly proportional to the frequency.
- ii. inversely proportional to the frequency.
- iii. proportional to the square of the frequency.
- iv. inversely proportional to the square of the frequency.

(b) The maximum current is
- i. linearly proportional to the frequency.
- ii. inversely proportional to the frequency.
- iii. proportional to the square of the frequency.
- iv. inversely proportional to the square of the frequency.

Section 22 - Resonance frequency in a series AC RLC circuit

22.1 The diagram shows an *RLC* circuit with two capacitors. If $R = 387\ \Omega$, $L = 3.65\times10^{-1}$ H, $C_1 = 1.12\times10^{-6}$ F and $C_2 = 3.27\times10^{-6}$ F, what is the resonance frequency of this circuit?

__________ rad/s

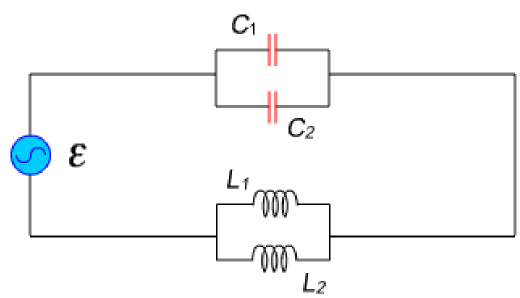


22.2 An *RLC* circuit contains a 4.10e−6 F capacitor in series with an inductor. What should the value of the inductance be for a resonance frequency of 63.0 rad/s?

__________ H

22.3 An AC circuit has a 1.00×10^{-11} F capacitor in parallel with a 3.00×10^{-11} F capacitor and a 6.00×10^{-2} H inductor in parallel with a 2.00×10^{-2} H inductor as shown in the picture. What is the resonant frequency of this circuit? Hint: Combine the inductors in the same manner as you would resistors.

__________ rad/s



22.4 The width of a resonance peak is often quantified using the width of the peak when the quantity is at half of the maximum height (this is called the "full width at half maximum" or FWHM). Find the FWHM of the current, in rad/s, in an *RLC* circuit in terms of the resistance *R*, the capacitance *C*, and the inductance *L*.

Section 23 - Rms power in AC resistor circuits

23.1 The emf of an AC circuit has an rms value of 120 V. (a) What is the maximum positive emf? (b) What is the most negative emf?

(a) __________ V
(b) __________ V

23.2 If the maximum current in an AC circuit is 9.00 A, what is the rms current?

__________ A

23.3 An AC current flows through a resistor, creating a maximum potential difference across it of 78.0 V. If the maximum current is 12.0 A, what is the average power dissipated by the resistor?

__________ W

23.4 The average power dissipated by a resistor in an AC circuit is 248 W. The rms current through the resistor is 15.6 A. What is the maximum potential difference across the resistor?

__________ V

Section 24 - Power in AC RLC circuits

24.1 What is the instantaneous power, at a time of 0.00350 seconds, dissipated by a 200 ohm resistor in a resonant AC circuit powered by a generator with a 21.0 V maximum potential difference operating at 5000 rad/s?

__________ W

24.2 A 275 Ω resistor is connected in series with a 3.45×10^{-2} H inductor and a 4.77×10^{-6} F capacitor, in an AC circuit driven by a generator with maximum emf 232 V and angular frequency 388 rad/s. What is the maximum power dissipated through the resistor?

__________ W

24.3 What is the average power dissipated by a 2550 Ω resistor in series with a 7.00×10^{-8} F capacitor and a 0.0200 H inductor when a generator with a maximum emf of 12.0 volts drives the circuit at 10000 rad/s.

__________ W

24.4 An electrical appliance draws an rms current of 10.0 A and has an average power of 720 W when connected to a power source with an rms emf of 120 V at 60.0 Hz. (a) What is the impedance of the appliance? The appliance could be analyzed as a series combination of a resistor, capacitor and inductor with the same electrical properties. (b) Considering the appliance in this fashion, what is the value of the series resistance? (c) The quantity $X_L - X_C$ (as found in the impedance equation) is called the *total reactance*. What is the magnitude of the total reactance of the appliance? (d) What is the instantaneous power at time t = 0 s?

(a) __________ Ω
(b) __________ Ω
(c) __________ Ω
(d) __________ W

Section 26 - Interactive problem: RLC radio tuner

26.1 Use the simulation in the interactive problem in this section to calculate the capacitance needed to tune the radio to the desired frequency.

__________ F

Additional Problems

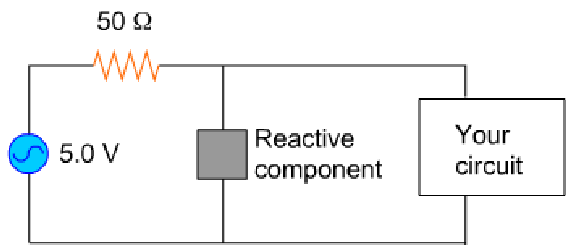
A.1 A series *RLC* circuit contains an AC generator (120 volts rms, frequency = 60 Hz), a 35.0 ohm resistor, a 1.49×10^{-5} F capacitor, and a 0.0300 H inductor. (a) What is the resonant frequency of the circuit? Express your answer in Hz. (b) Determine the capacitive reactance. (c) Determine the inductive reactance. (d) Determine the maximum current in the circuit. (e) Determine the phase constant. (f) Determine the average power. (g) What driving frequency minimizes the circuit impedance? (h) What driving frequency maximizes the average power consumed?

(a) __________ Hz
(b) __________ Ω
(c) __________ Ω
(d) __________ A
(e) __________ rad
(f) __________ W
(g) __________ Hz
(h) __________ Hz

A.2 One way to determine the inductance of an inductor is to place it in a series circuit with a resistor and capacitor of known values. The circuit is driven by a variable-frequency AC generator with a constant emf, and the frequency is varied until the current through the resistor, as measured by an AC ammeter, reaches a maximum. (a) A circuit like this is used with R = 235 Ω, $C = 1.50\times10^{-5}$ F, and maximum emf of 145 V. The current reaches a maximum when the frequency is 322 Hz. (a) What is the inductance? (b) What is the maximum current?

(a) __________ H
(b) __________ A

A.3 Suppose you have a circuit that can only operate properly at high frequencies. In order to guarantee correct operation you want to filter out any low frequency driving voltages. This is done using a *high pass filter*. Similarly, if you need a low frequency driving voltage you would want to use a *low pass filter*. A simple filter can be made with a single reactive component (a capacitor or inductor) which runs in parallel with your circuit as shown in the diagram (this way the voltage across the component is the same as the voltage across your circuit).



Consider a generator with a 5.0 V maximum potential difference in series with a 50 Ω resistor and a reactive component. If the generator is operating at 1.0 rad/s, what is the peak voltage across the reactive component if it is a (a) 1.0×10^{-6} F capacitor? (b) a 1.0×10^{-2} H inductor? Find the peak voltage across the reactive component if the generator operates at 1.0×10^{4} rad/s and the reactive component is a (c) 1.0×10^{-6} F capacitor. (d) a 1.0×10^{-2} H inductor. Do the same for a generator operating at 1.0×10^{9} rad/s with (e) a 1.0×10^{-6} F capacitor and (f) a 1.0×10^{-2} H inductor. Round your answers to the nearest one-hundredth of a volt. (g) Which component creates a low pass filter? (h) a high pass filter?

(a) __________ V

(b) __________ V

(c) __________ V

(d) __________ V

(e) __________ V

(f) __________ V

(g) i. The capacitor
ii. The inductor

(h) i. The capacitor
ii. The inductor

Electromagnetic Radiation

34.0 - Introduction

Electromagnetic radiation: Rainbows and radios. Sundazzled reflections. Shadowlamps and lampshadows. Red, white, and blue.

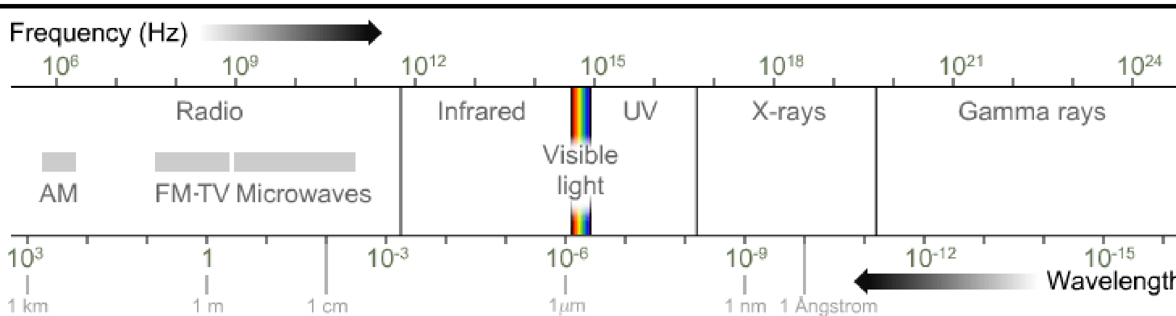
Radio and television signals, x-rays, microwaves: Each is a form of electromagnetic radiation. If steam and internal combustion engines symbolize the Industrial Revolution, and microprocessors and memory chips now power the Information Revolution, it almost seems that we have neglected to recognize the "Electromagnetic Revolution." Think about it: Can you imagine life without television sets or cell phones? You may long for such a life, or wonder how people ever survived without these devices!

These examples are from the world of engineered electromagnetic radiation. Even if you think we might all prosper without such technologies to entertain us, do our cooking, carry our messages, and diagnose our illnesses, you would be hard-pressed to survive without light. This form of electromagnetic radiation brings the Sun's energy to the Earth, warming the planet and supplying energy to plants, and in turn to creatures like us that depend on them. There are primitive forms of life that do not depend on the Sun's energy, but without light there would be no seeing, no room with a view, no sunsets, and no Rembrandts.

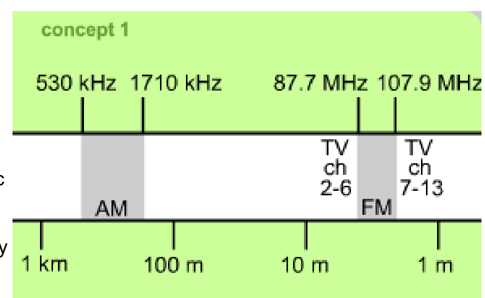
Some of the electromagnetic radiation that reaches your eyes was created mere nanoseconds earlier, like the light from a lamp. Other electromagnetic radiation is still propagating at its original speed through the cosmos, ten billion years or more after its birth. An example of this is the microwave background radiation, a pervasive remnant of the creation of the universe that is widely studied by astrophysicists.

Back here on Earth, this chapter covers the fundamental physical theory of electromagnetic radiation. Much of it builds on other topics, particularly the studies of waves, electric fields and magnetic fields.

34.1 - The electromagnetic spectrum



Electromagnetic spectrum: Electromagnetic radiation ordered by frequency or wavelength.



Electromagnetic spectrum
Radio waves in AM, FM bands

Electromagnetic radiation is a traveling wave that consists of electric and magnetic fields. Before delving into the details of such waves, we will discuss the electromagnetic spectrum, a system by which the types of electromagnetic radiation are classified.

The illustration of the electromagnetic spectrum above orders electromagnetic waves by frequency and by wavelength. In the diagram, frequency **increases** and wavelength **decreases** as you move from the left to the right. The chart's scale is based on powers of 10. Wavelengths range from more than 100 meters for AM radio signals to as small as 10^{-16} meters for gamma rays.

All electromagnetic waves travel at the same speed in a vacuum. This speed is designated by the letter c and is called the speed of light. (The letter c comes from *celeritas,* the Latin word for speed. It might be more accurate to refer to it as the speed of electromagnetic radiation.) The speed of light in a vacuum is exactly 299,792,458 m/s, and it is only slightly less in air.

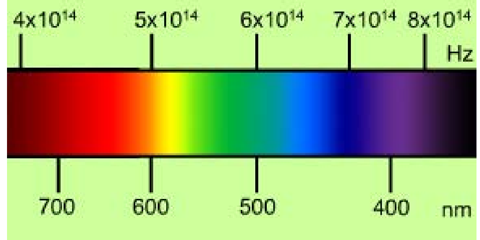
The unvarying nature of this speed has an important implication: The wavelength of electromagnetic radiation is inversely proportional to its frequency. As you may recall, the speed of a wave equals the product of its frequency and wavelength. This means that if you know the wavelength of the wave, you can determine its frequency (and vice versa). For instance, an electromagnetic wave with a wavelength of 300 meters, in the middle of the AM radio band, has a frequency of 1×10^6 Hz. This equals 3×10^8 m/s, the speed of light, divided by 300 m. The frequencies of electromagnetic waves range from less than one megahertz, or 10^6 Hz, for long radio waves to over 10^{24} Hz for gamma rays.

We will now review some of the bands of electromagnetic radiation and their manifestations. The lowest frequencies are often utilized for radio signals. AM and FM *radio waves* are typically produced by transmitters that incorporate electric oscillator circuits attached to antennas. The AM and FM radio bands are shown in Concept 1. *Microwave radiation* is at the upper end of the radio band, and is used for cellular telephone transmissions as well as for heating food in microwave ovens.

Infrared radiation has a higher frequency than microwaves and is associated with heat. It is generated by the thermal vibration or rotation of atoms and molecules.

Visible light is next as we go up the frequency range. It is of paramount importance to human beings, although it occupies only a small portion of the electromagnetic spectrum. Like some other forms of electromagnetic radiation, it is created when atoms emit radiation as their electrons drop from higher to lower energy levels. Light consists of electromagnetic waves that oscillate more than 100 trillion (10^{14}) times a second. The wavelengths of the various colors of light are in the hundreds of nanometers. Red light has the lowest frequency and longest wavelength, while violet light has the highest frequency and shortest wavelength. The visible light spectrum is shown in Concept 2.

concept 2



Visible light
Consists of spectrum of colors
· White = mixture of all colors
· Black = lack of visible light

The Sun emits a broad spectrum of electromagnetic radiation, including *ultraviolet* (UV) waves, with frequencies higher than those of visible light. These waves are the main cause of sunburn: Sunscreen lotion is designed to prevent them from reaching and harming your skin. This radiation can also harm your eyes, especially if you wear plastic sunglasses that diminish the amount of visible light reaching your eyes, but do not block UV rays. Typically, you squint when your eyes are exposed to strong light, and this helps protect them. If you wear sunglasses that do not stop ultraviolet light, your pupils will dilate, allowing an extra dose of harmful ultraviolet waves to enter your eyes.

The Earth's atmosphere, specifically the layer that contains a molecular form of oxygen called *ozone*, absorbs a great deal of the Sun's ultraviolet radiation, protecting plants and animals from its harmful effects. However, substances once commonly used in refrigerators and aerosols catalyze ozone-destroying chemical reactions in the atmosphere. Fortunately, the use of such substances has been restricted, but a "hole" in the ozone, a region where the amount of ozone has been significantly depleted, has been created above Antarctica. This hole varies in size from year to year, but on average is approximately the size of North America.

concept 3

X-rays
Used in
· Medical diagnosis and research
· X-ray diffraction studies
· X-ray astronomy

X-rays, the next band of electromagnetic frequencies, are even more dangerous than UV, but they are also useful. Doctors can use them to "see" shadowy images of the inside of the human body because they travel more easily through some tissues than others. Scientists also use them to discern the detailed crystalline structure of materials and to deduce the spatial configuration of complex molecules. NASA launched the Chandra X-ray Observatory in 1999 to capture the radiation emitted from high-energy regions of the universe, such as the parts of space around exploded stars. The acceleration of high-energy electrons is one source of x-rays. Because these rays can damage or destroy living cells and tissues, human exposure to them must be strictly limited.

Finally, *gamma rays* are the highest frequency electromagnetic waves. They are emitted by atomic nuclei undergoing certain nuclear reactions, as well as by high energy astronomical objects and events. These rays may enter the Earth's atmosphere from space, but they usually collide with air molecules and do not reach the ground. Those few that do reach the Earth can cause mutations in the DNA of living cells. Gamma radiation is highly destructive, which is one reason for the thick shielding used to protect workers from nuclear materials. However, like x-rays, gamma rays also have some beneficial uses. "Gamma knife" radiosurgery uses concentrated beams of gamma rays to kill cancer cells.

The distinction between the different parts of the spectrum appears clearer in the diagram above than it is in reality; the labeled bands actually blend into each other. For instance, there is no clear-cut frequency at which the shortest "radio waves" stop and the longest "far infrared" waves start; the bands in the diagram simply provide a convenient way to classify frequency ranges. Frequency classifications do become precise when business people enter the discussion: The rights to use certain frequency ranges, such as those for radio and television stations and cell phones, are worth tens of billions of dollars.

Despite the different names − AM radio signals, visible light, x-rays, gamma rays − the phenomena organized in the chart above and illustrated to the right are all forms of electromagnetic radiation. In the same way that you think of both tiny puddle ripples and long, slow ocean swells as being water waves, so you should think about the types of electromagnetic radiation. The frequency of the wave does not alter the fundamental laws of physics that govern it.

Electromagnetic wave: A wave consisting of electric and magnetic fields oscillating transversely to the direction of propagation.

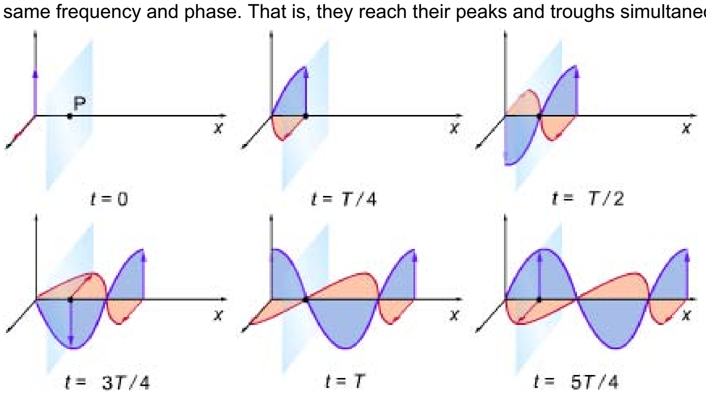
Physicist James Clerk Maxwell's brilliant studies pioneered research into the nature of electromagnetic waves. He correctly concluded that oscillating electric and magnetic fields can constitute a self-propagating wave that he called electromagnetic radiation. His law of induction (a changing electric field causes a magnetic field) combined with Faraday's law (a changing magnetic field causes an electric field) supplies the basis for understanding this kind of wave.

As the diagrams to the right show, the electric and magnetic fields in an electromagnetic wave are perpendicular to each other and to the direction of propagation of the wave. These illustrations also show the amplitudes of the fields varying sinusoidally as functions of position and time. Electromagnetic waves are an example of *transverse waves*. The fields can propagate outward from a source in all directions at the speed of light; for the sake of visual clarity, we have chosen to show them moving only along the x axis.

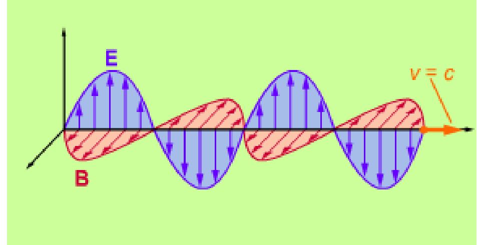
The animated diagram in Concept 2 and the illustrations below are used to emphasize three points. First, the depicted wave moves away from the source. For example, if you push the "transmit" button on a walkie-talkie, a wave is initiated that travels away from the walkie-talkie.

Second, at any fixed location in the path of the wave, both fields change over time. The wave below is drawn at intervals that are fractions $T/4$ of the period T. Look at the point P below, on the light blue vertical plane. The vectors from point P represent the direction and strength of the electric and magnetic fields at this point. As you can see, the vectors, and the fields they represent, change over time at P. Concept 2 shows them varying continuously with time at the point P.

Third, the diagrams reflect an important fact: The electric and magnetic fields have the same frequency and phase. That is, they reach their peaks and troughs simultaneously.

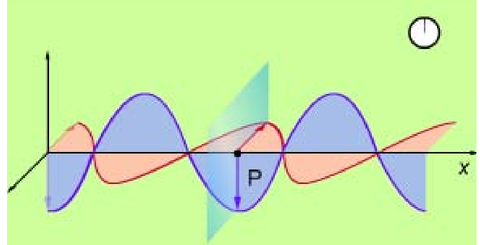


concept 1



Electromagnetic waves
Consist of electric and magnetic fields
· Perpendicular to each other
Propagate as transverse waves
· Perpendicular to direction of travel

concept 2



Electric, magnetic fields
Vary in strength over time at each point
Have same frequency and are in phase
Drive each other by changing strength

A wave on a string provides a good starting point for understanding electromagnetic waves. Both electromagnetic radiation and a wave on a string are transverse waves. The strengths of the two fields constituting the radiation can be described using sinusoidal functions, just as we can use a sinusoidal function to calculate the transverse displacement of a particle in a string through which a wave is moving.

There is a crucial difference, though: Electromagnetic radiation consists of electric and magnetic fields, and does not require a medium like a string for its propagation. Electromagnetic waves can travel in a vacuum. If this is troubling to you, you are in good company. It took some brilliant physicists a great deal of hard work to convince the world that light and other electromagnetic waves do not require a medium of transmission.

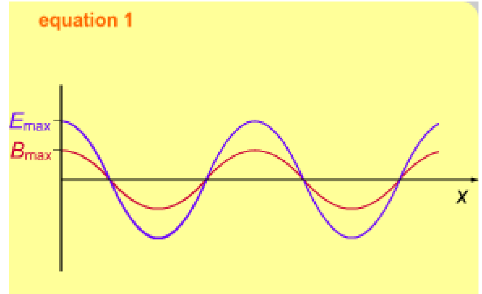
Furthermore, when electromagnetic waves radiate in all directions from a compact source like an antenna or a lamp, the radiation emitted at a particular instant travels outward on the surface of an expanding sphere, and its strength diminishes with distance from the source. The waves cannot be truly sinusoidal, since the amplitude of a sinusoidal function never diminishes.

In the sections that follow we will analyze *plane waves*, which propagate through space, say in the positive x direction, in parallel planar wave fronts rather than expanding spherical ones. They are good approximations to physical waves over small regions that are distant from the source of the waves. Plane waves never diminish in strength; they can be accurately modeled using sinusoidal functions, and we will do so.

34.3 - Proportionality of electric and magnetic fields

Although the electric and magnetic field vectors of an electromagnetic wave point in perpendicular directions, their magnitudes are strictly proportional to each other at all positions and at all times. We graphically display the magnitudes at a particular instant on the same coordinate system in Equation 1. Their proportionality is expressed in the equation to the right, using a constant c that depends on two other fundamental physical constants.

This proportionality turned out to have important implications in the study of electromagnetic radiation. Why? Because when calculated, the value of c was very close to the measured speed of light. This crucial discovery accelerated the understanding of the relationship between electromagnetic radiation such as light or radio waves, and electric and magnetic fields.



Proportionality of electric and magnetic field strengths

In an electromagnetic wave,

$$\frac{E}{B} = c$$

E = electric field strength

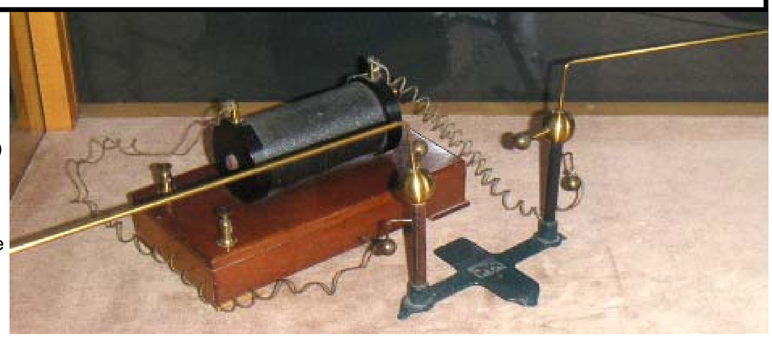
B = magnetic field strength

$c = 1/\sqrt{\mu_0 \varepsilon_0}$ = speed of wave

34.4 - Calculating the speed of light from fundamental constants

One of the major discoveries of 19th century physics was that light is a form of electromagnetic radiation. By applying and extending their knowledge of electric and magnetic fields, physicists were able both to create electromagnetic radiation (initially radio waves) and to predict the speed of the waves.

James Maxwell published his four laws governing electromagnetic phenomena in 1864, and at the same time he predicted the existence of self-propagating electromagnetic waves. His work enabled him to calculate what the speed of these waves would have to be.



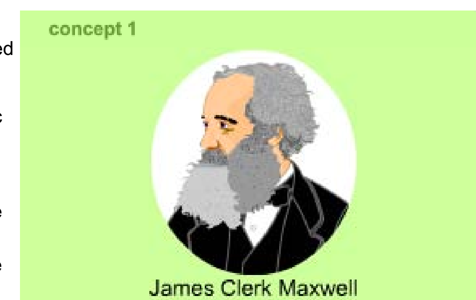
Heinrich Hertz transmitter, 1888. Transformer voltage causes a spark to jump between the postionable brass spheres, generating a radio pulse.

The angular frequency of any wave is $\omega = 2\pi f$, where f is its frequency in cycles per second. The angular wave number is $k = 2\pi/\lambda$, where λ is the wavelength. The speed of any wave is $v = \lambda f$. This means that the speed in terms of the angular frequency and wave number is $v = (2\pi/k)(\omega/2\pi)$, or ω/k.

Maxwell had already shown that the wave speed ω/k is a constant for electromagnetic waves, a constant he had designated as c and expressed in terms of μ_0 and ε_0. We state the relationship of c to these several variables and constants in Equation 1.

The fundamental constants μ_0 and ε_0 are used elsewhere in physics. For example, the permeability constant is used in equations that describe the magnetic field created by various electric current configurations. The permittivity constant is used to express one form of Coulomb's law. In other words, these constants determine the strengths of the electric and magnetic forces in the physical universe.

The example problem on the right asks you to repeat Maxwell's calculation of the value of c. The result is 2.998×10^8 m/s. In the year 1864, the speed of visible light in a vacuum had been known with fair accuracy for well over a century. The English physicist James Bradley estimated it in 1728 to be 3.1×10^8 m/s, based on his study of "stellar aberration," or the apparent change in the positions of stars as the Earth moves around the Sun. Because the calculated and observed speeds were so close, Maxwell's results provided the first evidence that light is a kind of electromagnetic wave.



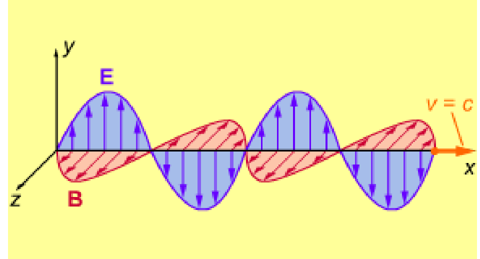
Speed of light

Theoretical speed of radiation
= the measured speed of light

Conclusion: Light is electromagnetic radiation!

Understanding that light is an electromagnetic wave, and knowing the general relationship between frequency and wavelength, sparked the discovery of additional types of electromagnetic radiation. In 1888 Heinrich Hertz created what we would now call radio receivers and transmitters, one of which you see in the illustration above. He proved the existence, and wavelike nature, of radiation having frequencies around 100 MHz.

equation 1

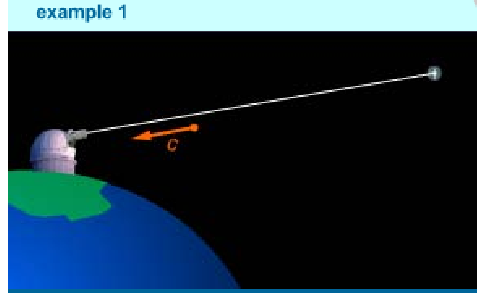


Speed of an electromagnetic wave

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

c = speed of electromagnetic wave
ω = angular frequency of wave
k = angular wave number of wave
μ_0 = permeability of free space
Constant $\mu_0 = 4\pi\times10^{-7}$ T·m/A
ε_0 = permittivity of free space
Constant $\varepsilon_0 = 8.854\times10^{-12}$ C²/N·m²

example 1



What is the speed of an electromagnetic wave in a vacuum?

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

$\mu_0\varepsilon_0 = (4\pi\times10^{-7})(8.854\times10^{-12})$
$\mu_0\varepsilon_0 = 1.113\times10^{-17}$ s²/m²
$\sqrt{\mu_0\varepsilon_0} = 3.336\times10^{-9}$ s/m
$c = 2.998\times10^{8}$ m/s

34.5 - Creating electromagnetic waves: antennas

Radio antennas create electromagnetic waves. A radio antenna is part of an overall system called a radio transmitter that converts the information contained in sound waves into electromagnetic waves. A radio receiver then reverses the process, converting the signals from electromagnetic waves back to sound waves.

Radio-wave transmitters. A cable carries a modulated electric signal to the dipole antenna rod assembly visible at the top of each tower.

The system depicted to the right shows the fundamentals of a radio transmitter. In the illustrations, the terminals of an AC generator are connected to two rods of conducting material: an antenna. The AC generator produces an emf $\mathcal{E}$ that varies sinusoidally over time. The emf drives positive and negative charges to opposite ends of the antenna. The **separation** of the charges on the rods produces an electric field. The AC generator causes the amount and sign of the charge on each rod to vary over time, so that the resulting electric field varies in strength and orientation as well. (The flow of charge − that is, the current − also produces a varying magnetic field close to the antenna, part of what is called the *near field*, which we do not show here.)

The electric field produced by the antenna at each instant in time propagates outward in all directions at the speed of light. For simplicity's sake, we only show it traveling in the positive x direction in Concept 2. The electric field changes continuously with time and that change induces a magnetic field. To be precise, it induces a magnetic field proportional to the rate of change of the electric flux with respect to time, as described by Maxwell's law of induction. In turn, the changing magnetic field regenerates the electric field as the wave travels.

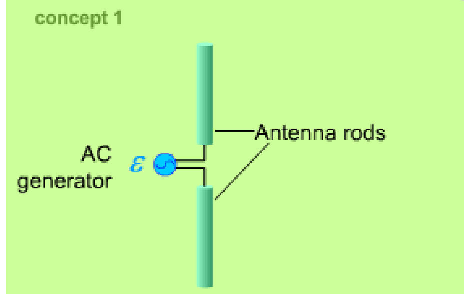
This coupling of changes in the magnetic and electric fields enables the electromagnetic wave to cross vast gulfs of space over immense spans of time. Electromagnetic radiation from distant stars, including light, reaches the Earth after billions of years of travel.

How does an antenna differ from other circuits you may have studied in which current flows or charge is stored? Consider a battery-resistor circuit or a battery-capacitor circuit in equilibrium; the current in the first creates a constant magnetic field, and the stored charge in the second creates a constant electric field. Both fields rapidly diminish as they extend outward in space. The crucial difference with the antenna is that not only does charge accumulate at its ends, but the AC generator continually causes the distribution of charge to change. The electric field varies sinusoidally over time, and a constantly changing electric field is the crucial element required to create continuous, self-propagating electromagnetic radiation.

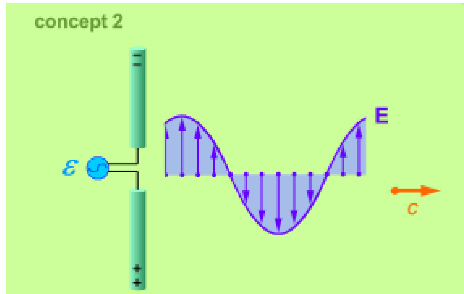
Electromagnetic waves are generated when charges move at nonconstant velocities, as in an antenna. That is, they are generated by **accelerating** charges. In an antenna, the acceleration is in a straight line. Charged particles moving in uniform circular motion also emit electromagnetic radiation, called *synchrotron radiation*, due to their centripetal acceleration.

The AM and FM radio bands are located in different parts of the electromagnetic spectrum, and they are used in different ways to broadcast program content. The difference between them consists in how the information they convey is encoded. In *amplitude modulated* or *AM radio*, sound waves are encoded by varying the amplitude of a *carrier radio wave* around some reference value. Changes in amplitude convey the signal. The frequency of the carrier wave is around 1 MHz for AM radio.

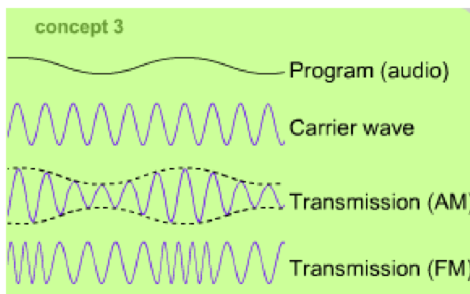
In *frequency modulated* or *FM radio*, sound is encoded by slightly varying the frequency of the carrier wave around its base frequency. For FM radio, and television, the carrier wave frequency extends upward from around 100 MHz.



Simple antenna components
A source of varying emf
Two conducting rods



Simple antenna operation
Generates varying charge on rods
Charged rods create electric field
Electric field varies over time
Changing E-field induces B-field



AM and FM radio
Radio wave is "carrier"
AM varies amplitude of carrier
FM varies frequency of carrier

34.6 - Electromagnetic energy: the Poynting vector

Electromagnetic waves transmit energy that is crucial to life on Earth. The process of photosynthesis turns the energy of sunlight into chemical energy used by plants, and by the animals that subsist on them − or on each other.

Electromagnetic energy from the Sun powers photosynthesis.

This section introduces the *Poynting vector*, which begins our discussion of the energy transported by an electromagnetic wave. It will prove a useful tool in developing a formula for the *intensity* of electromagnetic radiation in the next section. The Poynting vector is represented with the letter $\mathbf{S}$; Equation 1 shows you how to calculate the Poynting vector of a wave in terms of its electric and magnetic fields.

The diagram of Equation 1 shows an electromagnetic wave passing through a surface. The instantaneous rate at which energy is transported through the surface by the wave, per unit area, is called the *area power density* of the wave. The area power density is equal to the magnitude S of the Poynting vector.

The surface area through which the instantaneous power density is measured is perpendicular to the direction of the wave's propagation. When radiation reaches a physical surface obliquely, the cosine of its angle with the area vector can be used to calculate the power conveyed to the surface. This is analogous to the calculation of electric or magnetic flux.

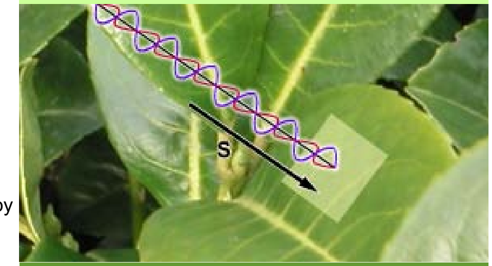
As Equation 1 shows, the Poynting vector equals the cross product of the vectors representing the electric and magnetic fields of the electromagnetic radiation, divided by the permeability constant. Since these fields are always perpendicular to one another, the sine of the angle between them, used to evaluate the magnitude of the cross product, always equals one, and can be effectively ignored when calculating the instantaneous area power density S. The units of the Poynting vector are watts per square meter.

The direction of $\mathbf{S}$ is determined by the right-hand rule. If you apply the rule, wrapping your fingers from $\mathbf{E}$ to $\mathbf{B}$ and noting the direction of your thumb, you can correctly determine that it is parallel to the direction of propagation of the wave. When $\mathbf{E}$ reverses its direction, so does $\mathbf{B}$, and the direction of $\mathbf{S}$ remains the same, "pointing" (heh, heh) in the direction of the wave's motion.

As an electromagnetic wave passes through a surface, the strengths of its electric and magnetic fields there change sinusoidally with time. Since the Poynting vector is the product of these fields, it changes sinusoidally over time, as well. In fact, it varies with values between zero and $E_{\max}B_{\max}/\mu_0$, with a frequency twice that of the fields.

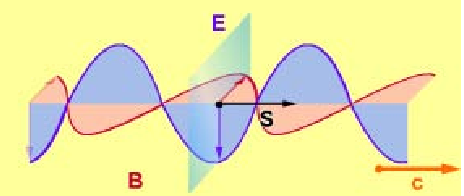
If you are curious why it has this frequency, recall from the field equations that E and B are both cosine functions of time at a fixed point. Then use the trigonometric identity $\cos^2 t = [1 + \cos 2t]/2$.

concept 1



Poynting vector
Power per unit surface area
Surface perpendicular to wave direction

equation 1



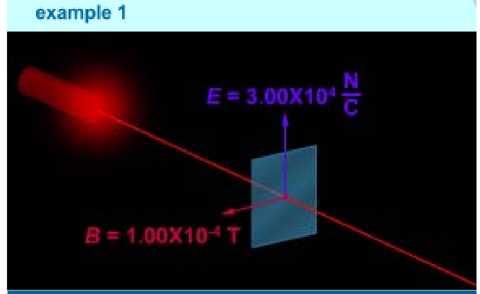
Poynting vector

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}$$

$\mathbf{S}$ = Poynting vector
S = instantaneous area power density
μ_0 = permeability of free space
$\mathbf{E}$ = electric field
$\mathbf{B}$ = magnetic field

Units: watts per square meter (W/m^2)

example 1



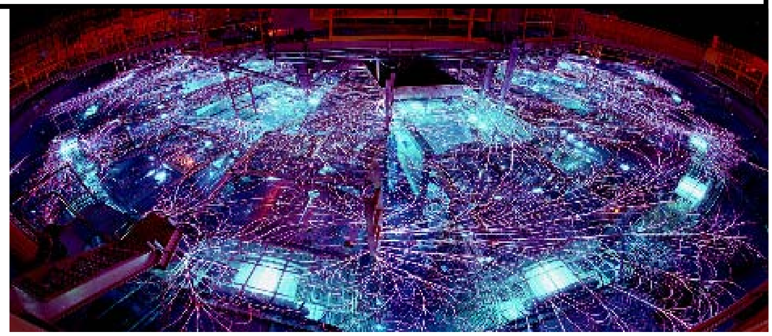
At this instant, what is the area power density of the ruby laser light?

$S = EB / \mu_0$

$$S = \frac{(3.00\times10^{4}\,\frac{\text{N}}{\text{C}})(1.00\times10^{-4}\,\text{T})}{4\pi\times10^{-7}\,\frac{\text{T}\cdot\text{m}}{\text{A}}}$$

$S = 2.39\times10^{6}$ W/m^2

Intensity: Average rate of power transmission per unit area.



The Z Accelerator uses extremely high energy densities to produce fusion.

As you have seen, the instantaneous area power density of an electromagnetic wave equals the magnitude S of the Poynting vector. It is proportional to the product of the strengths of the electric and magnetic fields at any instant in time.

This section focuses on two other properties used to describe the power or energy of a wave: the intensity and the average volume energy density, which are time-average values rather than instantaneous ones.

Why are average values useful? One reason is that the instantaneous area power density of an electromagnetic wave fluctuates very rapidly. For example, for green light it oscillates at 10^{15} Hz. Most radiation detectors, including the human eye, cannot distinguish between such a rapid oscillation and a constant power density. The time average of the area power density is called the **intensity** of a wave. Intensity is, roughly speaking, what you perceive as the "brightness" of a wave of visible light. It is analogous to the intensity of a sound wave, and its units are the same: W/m^2.

As already stated, the magnitude of the Poynting vector describes instantaneous power density and it is a good starting point for determining intensity: This magnitude equals EB/μ_0. Using the relationship $E/B = c$, we can restate the instantaneous power density as $E^2/\mu_0 c$, where the electric field strength E varies as a sinusoidal function of time.

The average value of any squared sinusoidal function is one half the squared amplitude of the function. Here, this means $(E^2)_{avg} = E^2_{max}/2$. Substituting this time-average value into the equation $S = E^2/\mu_0 c$ gives the formula for intensity, symbolized by the letter I, that appears in Equation 1.

With oscillating quantities, scientists often prefer to express an average value in terms of the *root mean square*. For a sinusoidal function, this equals the maximum value of the function divided by the square root of two. We express intensity in terms of the root mean square of the electric field with the second boxed formula in Equation 1, which follows from the first boxed equation and the definition of the root mean square.

When scientists measure the "power density" of an electromagnetic wave, they measure the amount of energy transported by the wave per unit time through a unit of surface area. An alternate approach to density is to select a fixed instant in time and measure the amount of electric and magnetic energy contained in a volume of space at that instant. When scientists take this approach, they are talking about *energy density*, or volume energy density. The units of energy density are joules per cubic meter, J/m^3.

In Equations 2 and 3 on the right, you see equations for energy density. The energy density u_E of an electric field $\mathbf{E}$ is $\varepsilon_0 E^2/2$, a result derived in the study of capacitors. The energy density u_B of a magnetic field $\mathbf{B}$ is $B^2/2\mu_0$, a relationship derived in the study of inductors. The formulas hold true, not just for the fields in circuit components, but for any electric and magnetic fields. Since these fields are constantly changing in an electromagnetic wave, the equations reflect **instantaneous** energy densities.

It can be shown, using the relationships $B = E/c$ and $c^2 = 1/\mu_0\varepsilon_0$, that even as they vary over time, the electric and magnetic energy densities are equal to each other at all times and all positions in an electromagnetic wave. You see this equality stated in Equation 2. From the standpoint of energy, electromagnetic radiation is "half electric" and "half magnetic." The **total** energy density u is defined as the sum of these two equal quantities. This is shown as the final equation in Equation 2.

In a fashion similar to the one used to calculate intensity, a formula for the average over time of the total energy density in terms of E_{max} can be derived. It is shown in Equation 3.

The "Z Accelerator" in Albuquerque, New Mexico (pictured above) has achieved the controlled fusion of a BB-sized capsule of deuterium ("heavy" hydrogen whose atoms have a neutron plus a proton in the nucleus) by bombarding it with 20 trillion watts of x-radiation for the brief period of five nanoseconds. The average total energy density of this bombardment is 3×10^{12} J/m^3. The hydrogen in the sample fuses to form helium and produces enough energy to power a dim light bulb for one ten-thousandth of a second. Someday, such fusion reactors may provide a practical source of electric power.

The first example problem on the right poses a more mundane challenge. It asks: What are the maximum electric and magnetic field strengths of white light emanating from a computer screen? The intensity of the light is given.

equation 1

Intensity

Average value of S over time

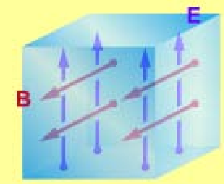
$$I = S_{avg} = \frac{E^2_{max}}{2\mu_0 c}$$

I = wave intensity
S_{avg} = average area power density
E_{max} = electric field amplitude

$$I = \frac{E^2_{rms}}{\mu_0 c}$$

E_{rms} = root mean square electric field
Units: watts per square meter (W/m^2)

equation 2



Electric, magnetic energy densities

$$u_E = \frac{\varepsilon_0 E^2}{2} \qquad u_B = \frac{B^2}{2\mu_0}$$

Since $E^2/B^2 = c^2 = 1/\mu_0\varepsilon_0$, then

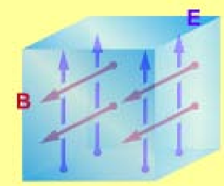
$$u_E = u_B$$

The total energy density is

$$u = u_E + u_B = 2u_E = 2u_B$$

u_E = electric field energy density
u_B = magnetic field energy density
u = total energy density

equation 3

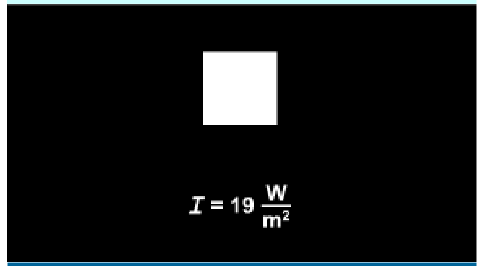


Average energy per unit volume
Average value of u over time

$$u_{avg} = \frac{\varepsilon_0 E_{max}^2}{2}$$

u_{avg} = average total energy density
Units: joules per cubic meter (J/m^3)

example 1



What are the maximum field strengths E_{max} and B_{max} for the light emanating from the white square above?

$$I = \frac{E_{max}^2}{2\mu_0 c}\text{, so } E_{max} = \sqrt{2\mu_0 cI}$$

$$E_{max} = \sqrt{2(4\pi \times 10^{-7})(3.0 \times 10^{8})(19)}$$

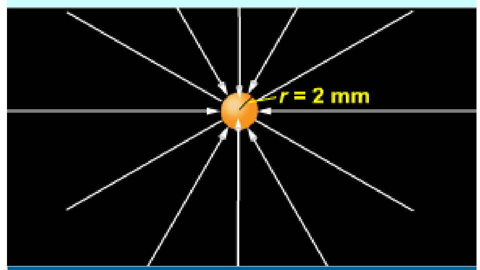
E_{max} = 120 V/m

$B_{max} = E_{max}/c$

B_{max} = (120 V/m)/(3.0×10^8 m/s)

B_{max} = 4.0×10^{-7} T

example 2



A fusion reactor focuses 2×10^{13} watts of x-rays on this deuterium capsule for 5×10^{-9} s. What is the average energy density of the radiation? What is the electric field amplitude of the radiation?

$$u_{avg} = \frac{E}{Vol} = \frac{P \cdot t}{Vol}$$

$$u_{avg} = \frac{(2 \times 10^{13}\ \text{W})(5 \times 10^{-9}\ \text{s})}{\frac{4}{3}\pi (0.002)^3\ \text{m}^3}$$

$$u_{avg} = 3 \times 10^{12}\ \text{J/m}^3$$

$$u_{avg} = \frac{\varepsilon_0 E_{max}^2}{2}\text{, so } E_{max} = \sqrt{\frac{2u_{avg}}{\varepsilon_0}}$$

$$E_{max} = \sqrt{\frac{2(3 \times 10^{12}\ \frac{\text{J}}{\text{m}^3})}{8.85 \times 10^{-12}\ \frac{\text{C}^2}{\text{N} \cdot \text{m}^2}}} = 8 \times 10^{11}\ \frac{\text{V}}{\text{m}}$$

34.8 - Earth's seasons

Almost all the energy we use on Earth originates in the Sun and arrives in the form of electromagnetic radiation. Roughly the same amount of solar power reaches the planet throughout the year, yet many places on the globe experience significant seasonal variations in the rate at which they receive this energy.

The cause of seasonal changes in the Earth's climate is the tilt of its axis, the line about which it rotates. The illustration in Concept 1 shows the position of the Earth in its orbit at different times of the year, as well as the direction in which the axis points. The axis is tilted at a $23.5°$ angle away from a line perpendicular to the Earth's orbital plane. It always points towards the same direction in space (which is why Polaris remains the North Star throughout the year).

March 21 and September 22 are known as the *equinoxes*. The name refers to the equal lengths of night and day (12 hours each) for all locations on Earth on these dates. December 21 and June 21 are the *solstices*. As the season progresses from autumn to winter, the Sun rises to a lower high point in the sky each day, and the days get shorter. On the winter solstice (meaning "sun stop"), the Sun stops getting lower and begins to rise to a higher apex each day, as it does through the rest of the winter and spring. The opposite happens after the summer solstice − the Sun once again peaks at a lower point each day. (The dates of the equinoxes and solstices vary from year to year, but are always around the 21st of the month.)

While June 21 is the summer solstice in the Northern Hemisphere, it is the winter solstice for the Southern Hemisphere. Concept 2 illustrates why this is true. It shows a Northern Hemisphere city, Beijing, and a Southern Hemisphere city, Perth, at noon on June 21.

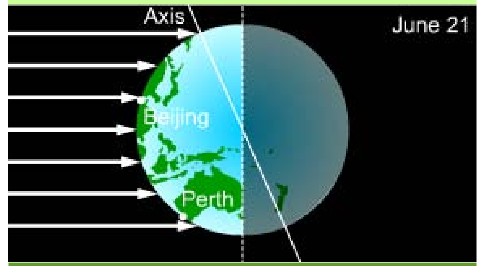
Imagine a solar collection plate of area A lying flat on the ground, tangent to the Earth's surface in either of these cities. Light rays from the Sun arrive approximately parallel to the plane of the Earth's orbit. On June 21, the sunlight intersects the collecting plate in Beijing at a steeper angle than in Perth. Because Beijing is receiving sunlight more vertically, the energy from that light is more concentrated − Beijing is receiving more power over the area of its collecting plate.

In Perth, a smaller amount of sunlight is being spread over the same collecting area because of the oblique, slanting angle at which it hits the plate. The plate absorbs less power. It is summer in Beijing, and winter in Perth.

concept 1

Cause of varying intensity
Earth's axis of rotation is tilted

concept 2



The seasons
Greater intensity in summer, less in winter
Northern, Southern Hemispheres have opposite seasons

You can also look at the situation from a flux perspective. In Beijing the sunlight is closer to being parallel to the area vector of the plate (considered to be pointing into the Earth), meaning there is a greater flux of light through the plate in Beijing. The greater flux means that more solar power is being received there.

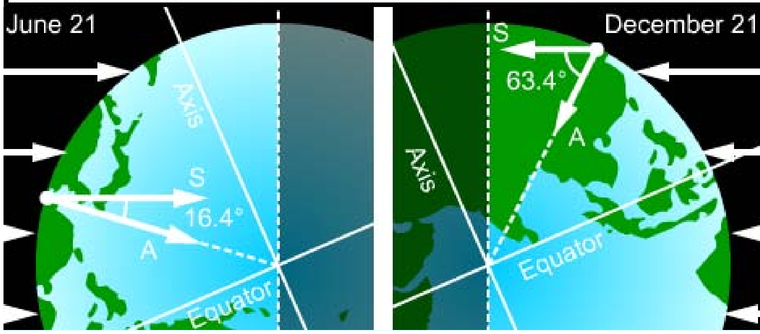
Six months later, on December 21, the situation will be reversed: Perth will receive more direct sunlight than Beijing.

Generally speaking, locations farther than Beijing or Perth from the equator experience a greater variation in the power they receive throughout the year, and places closer to the equator experience less change. At the poles − dark six months of the year − this difference is extreme.

Some people mistakenly believe that the seasons are due to the eccentricity of the Earth's orbit − the fact that the Earth's distance from the Sun changes throughout the year. Your first clue that this belief is false is the observation that summer in the Southern Hemisphere occurs at the same time as winter in the Northern Hemisphere (you merely have to make a long-distance phone call to confirm this). In fact, during winter in the Northern Hemisphere, the Earth is actually **closer** to the Sun than in summer. The reason the eccentricity has only a slight effect is that the Earth's orbit is only slightly elliptical. The annual variation in insolation due to the eccentricity of the Earth's orbit is about 7%, in contrast to an approximately 110% increase from winter to summer (at the latitude of Beijing) due to axial tilt.

In the next section we calculate the change in flux due to axial tilt in Beijing.

34.9 - Sample problem: seasonal flux



In Beijing, what is the ratio of solar flux on June 21 to the flux on December 21?

Assume that sunlight arrives parallel to the plane of the Earth's orbit and that its intensity has a constant value throughout the year.

To quantify the amount of sunlight, use the Poynting vector **S**, which equals the light's power per unit area. In the case of electric or magnetic fields, flux has been defined as the magnitude of a field vector times a component of an area vector. Here, **S** fulfills the role of the field vector and the dimensions of solar flux are "power received": power per unit area, times area, equals power. Beijing is at a latitude of $39.9°$, so the angle the collecting plate's area vector makes with the Poynting vector on June 21 is $39.9° - 23.5° = 16.4°$. On December 21 it is $39.9° + 23.5° = 63.4°$.

Variables

instantaneous power per unit area reaching Earth	$\mathbf{S}$
area vector of Earth-horizontal collecting plate	$\mathbf{A}$
angle between Poynting and area vectors	θ
solar flux on June 21	Φ_{Jun}
solar flux on December 21	Φ_{Dec}

What is the strategy?

1. Calculate the solar flux on June 21 as a multiple of the magnitudes of $\mathbf{S}$ and $\mathbf{A}$.
2. Calculate the solar flux on December 21 as a multiple of the magnitudes of $\mathbf{S}$ and $\mathbf{A}$.
3. Divide to obtain the ratio of solar fluxes.

Physics principles and equations

Equation of flux

$$\Phi = \mathbf{S} \cdot \mathbf{A} = SA \cos \theta$$

Step-by-step solution

Start by calculating the flux on June 21. Use the illustration.

Step	Reason
1. $\Phi_{\text{Jun}} = SA \cos \theta$	definition of flux
2. $\Phi_{\text{Jun}} = SA \cos 16.4°$	angle between Poynting and area vectors
3. $\Phi_{\text{Jun}} = 0.96\ SA$	evaluate cosine

Now calculate the flux on December 21.

Step	Reason
4. $\Phi_{\text{Dec}} = SA \cos \theta$	definition of flux
5. $\Phi_{\text{Dec}} = SA \cos 63.4°$	angle between Poynting and area vectors
6. $\Phi_{\text{Dec}} = 0.45\ SA$	evaluate cosine

Find the ratio of the flux in June to that in December.

Step	Reason
7. $\dfrac{\Phi_{\text{Jun}}}{\Phi_{\text{Dec}}} = \dfrac{0.96SA}{0.45SA}$ $\dfrac{\Phi_{\text{Jun}}}{\Phi_{\text{Dec}}} = 2.1$	calculate ratio

The ratio of sunlight power received in June versus December in Beijing is about 2.1:1. This means the increase in flux from winter to summer is 110%. Greater ratios hold for locations farther than Beijing from the equator, while ratios approaching 1.0 hold for places closer to the equator.

As stated before, this change is much greater than the 7% worldwide variation in flux due to the Earth approaching and receding from the Sun in the course of its yearly orbit.

34.10 - Radiation intensity and distance

So far, we have modeled electromagnetic waves as plane waves, whose average power is the same everywhere in space.

Now we want to consider *spherical waves*, waves emanating from a point source and propagating outward in all directions. On the right you see a spherical wave emanating from a point source and passing through a fixed spherical surface of radius *r*. With this kind of wave, the intensity of the radiation decreases as the distance from the source of the radiation increases. We will assume that the energy of the radiation is emitted from the source at a constant rate: The power of the source does not change.

To understand how intensity varies with distance, first consider how a spherical wave spreads. It travels in all directions at the same rate (the speed of light, for an electromagnetic wave). At any instant in time, every point on a wave front is the same distance from the source of the wave. This means the wave's energy is distributed over a spherical surface, since all the points on a sphere are the same distance from its center.

The propagation of spherical wave fronts is more or less analogous to what happens when a pebble is dropped into a still pond: Wave fronts move out in concentric circles from where the pebble strikes the water. However, spherical electromagnetic wave fronts spread out in three dimensions, not two.

To discuss intensity, say the source emits electromagnetic radiation with power P and consider the radiation passing out through a fixed sphere of radius r. The energy emitted at the source during a time interval Δt will pass through the whole surface of this sphere during a later time interval of equal duration, so the power is the same at the source as at the surface. Intensity equals average power per unit area. Using the fact that the area of the sphere equals $4\pi r^2$, we can state the intensity equation in Equation 1. The equation shows that the intensity of radiation from a point source obeys an inverse square law: It diminishes as the square of the distance from the source. (A similar equation describes the intensity of spherically expanding sound waves.)

Example 1 asks you to calculate the intensity of sunlight at the distance of the Earth's orbit. At this distance, the Sun can be considered a point source of electromagnetic radiation. The Sun actually emits radiation of varying intensity, mostly as visible light, over a variety of wavelengths. The result stated is the total intensity for all wavelengths of visible light: roughly speaking, what we perceive as "white" light.

The intensity of sunlight falling on the Earth's atmosphere is about 70 times the intensity of the white light emerging from your computer screen, which is around 19 W/m^2. The intensity of sunlight reaching the Earth's surface is less, due to factors such as the reflection, absorption and scattering of the light in the atmosphere.

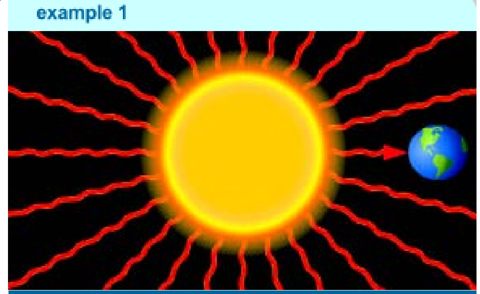
equation 1

Intensity and distance

$$I = \frac{P}{4\pi r^2}$$

I = radiation intensity
P = power of radiation source
r = distance from radiation source

example 1



The Sun emits 3.91×10^{26} watts of radiation. What is its intensity at the distance of the Earth, 1.50×10^{11} m?

$I = P / 4\pi r^2$
$I = (3.91\times10^{26}\text{ W}) / 4\pi(1.50\times10^{11}\text{ m})^2$
$I = 1380\text{ W/m}^2$

34.11 - Interactive checkpoint: an illuminated manuscript

The desk lamp in the picture uses a 100 W bulb, producing white light with an efficiency of 6.50%. The part of the book directly underneath, where the intensity of the light is greatest, is 25.0 cm from the bulb.
What is the amplitude of the electric field component of the light falling on the indicated portion of the book's page?

Answer:

E_{max} = ______ N/C

34.12 - Intensity and field strength around a dipole antenna

Radio wave antennas produce electromagnetic radiation by accelerating charges back and forth along a conductor. The waves, which travel outward from the antenna in many directions, constitute the *radiation field* of the antenna.

Antennas produce two kinds of fields: In addition to the radiation field, they produce a *near field* that is fundamentally different from the radiation field, and of less importance since its strength diminishes rapidly with distance from the antenna. In this section we discuss the differences between these fields and show how the radiation field expands not just along the x axis, but throughout three-dimensional space in a pattern of varying intensities.

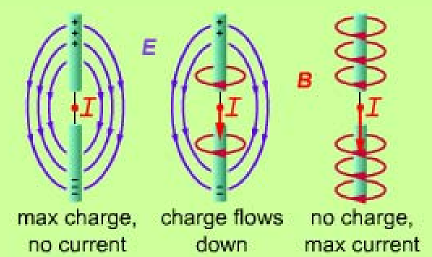
First, we discuss the near field, which like the radiation field has electric and magnetic components. Concept 1 shows a dipole antenna at three stages in its operation. In the first stage, the separation of charges on the antenna is at its maximum, and no current is flowing. At this point, the three-dimensional electric field around the antenna closely resembles that of an ideal electric dipole. This means its strength diminishes as the **cube** of the distance r from the center of the dipole.

In the second stage of operation, charge is starting to flow from the top to the bottom of the antenna. The electric field is diminishing in strength, and the current is beginning to create a circular magnetic field around the antenna. The magnetic field's direction is given by the right-hand rule for currents.

In the third stage, the charge separation is zero, so there is no electric field, but the current flow is at a maximum, generating the maximum magnetic field. The current will continue to flow downward in the antenna until it builds up a maximum separation of charge in the opposite direction, when once again the electric field will be at a maximum (directed upward), and the magnetic field will be zero.

By inspecting the diagram, you can see that the electric and magnetic components of the near field oscillate (they vary sinusoidally) and are perpendicular to each other, but they do not together constitute the electric and magnetic components of a self-propagating wave. This is because they are out of phase: The magnetic field lags the electric field by $90°$. Since the strength of the near field decreases as $1/r^3$, and it cannot propagate itself, at moderate distances from the antenna this field is barely detectable.

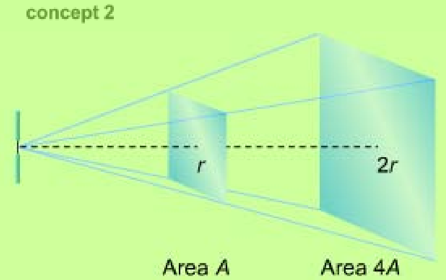
concept 1



The near field

Falls off as $1/r^3$
· E and B components 90° out of phase
· No self-propagation

concept 2

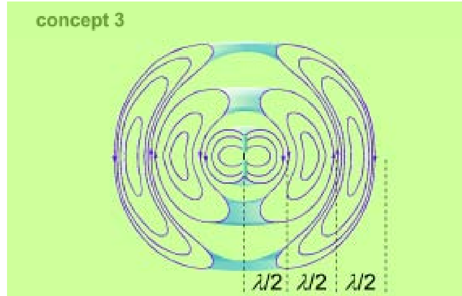


The radiation field

Falls off as $1/r$
· E and B components in phase
· Self-propagating waves

The self-propagating electromagnetic radiation field of the antenna arises from the sinusoidal variation within the electric and magnetic fields discussed above, rather than from their quite limited extent or from their nonexistent interaction with each other. And the radiation field decreases in strength much more gradually. Concept 2 shows a portion of the radiation field emerging from an antenna. Its intensity is measured in the equatorial plane of the dipole at two distances that are large enough that the antenna approximates a point source. As shown in a previous section, the intensity of such radiation decreases with increasing distance from the source: It is proportional to $1/r^2$. For example, at twice the distance from the antenna, the power of the wave is spread out over an area four times as large. Since intensity is proportional to the square of the amplitude of the electric field, the strength of the electric field itself must be decreasing as $1/r$.

In our previous discussion of antennas, for simplicity's sake we showed a single electromagnetic wave propagating along the positive x axis. However, waves actually propagate in many directions from an antenna. The illustration in Concept 3 shows the complex pattern of electric fields around a dipole antenna at an instant in time. The illustrated electric component of the radiation field consists of an expanding series of concentric tori, or "donuts," circling the equator of the dipole. The donuts farther away from the antenna were generated by charges on the antenna at earlier times. The field orientations in the donuts alternate due to the alternating polarity of the antenna charge separation. The field lines of the magnetic field component of the radiation field, which are not shown, would consist of concentric rings around the antenna, everywhere perpendicular to the electric field lines, and periodically alternating in orientation.



Structure of the radiation field
Series of expanding "donuts"
Maximum field strength on dipole equator
Zero field strength along dipole axis

An ideal source that generates a spherical radiation field is called an *isotropic radiator*. We used such an ideal radiator when we developed the formula expressing intensity as a function of the distance from a source of radiation. Concept 3 shows that the radiation field generated by a dipole antenna does not propagate in a series of expanding spheres. It is strongest where the field lines are closest together, that is, around the equator of the dipole. It is weakest along the axis of the dipole, where it has zero strength.

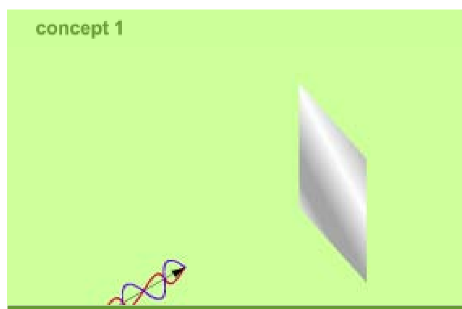
The dipole antennas we have been discussing are sometimes called *half-wave* dipole antennas. This is because they efficiently emit radiation with a wavelength equal to twice the combined length of the two antenna rods. The frequency of the electromagnetic wave is the same as the frequency of the alternating current driving the antenna. When the length of the antenna is half a wavelength, the wave's frequency is in resonance with the natural oscillatory motion of electric charges in the antenna, resulting in efficient operation.

34.13 - Radiation pressure

Interstellar sailing in "The Lady Who Sailed *The Soul*" by Cordwainer Smith.

Science fiction writers like to ponder space travel. Here is one of their creative ideas: giant space sails that use light instead of wind to move spacecraft. Radiation from the Sun or other stars exerts a force on the spaceship's sail, propelling the craft through space.

This concept has moved beyond science fiction. The 1964 Echo 2 satellite, a highly reflective low-mass Mylar balloon with a diameter of 30 m, experienced detectable acceleration as a result of sunlight pressure. A group called the Planetary Society is planning to launch a spacecraft that will use radiation pressure as its motive force.



Radiation pressure
Radiation striking surface transfers momentum
· It exerts force, pressure

The proposed propulsion of spacecraft depends on the fact that light has linear momentum. This assertion would have come as no surprise to Isaac Newton, who is almost as famous for his use of a prism to decompose white light into its constituent colors as he is for his studies of gravity and motion. He theorized that light consisted of streams of tiny particles that he called "corpuscles," and was thereby able to explain many of its observed properties.

According to a particle theory of light, a stream of light corpuscles striking a surface imparts momentum to it, much as a hail of bullets might cause a target to move. This effect is shown in Concept 2.

During the 17th century, Newton's particle theory of light vied with a wave theory championed by his rivals Christiaan Huygens and Robert Hooke. For over a hundred years Newton's corpuscles held sway, until Thomas Young's 1801 diffraction experiments (the topic of another chapter in this book) made the wavelike nature of light incontestable.

Maxwell, with his famous treatise on the nature and speed of electromagnetic waves, and his apparent discovery that light itself is a form of electromagnetic radiation, seemed to drive the final nail into the coffin of the particle theory. He was able to explain radiation pressure by showing that electromagnetic waves would possess momentum exactly as particles did. The debate did not end there, however. Albert Einstein used a phenomenon called the photoelectric effect to demonstrate the particle nature of light. Today, light is understood to exhibit both wavelike and particlelike properties.

To show how much pressure radiation exerts on an object, we start with an equation developed by Maxwell that is shown in Equation 1. He determined that the amount of momentum transferred by radiation to a "blackbody" equals the amount of energy it transfers to it, divided by the speed of light. A *blackbody* is an object that absorbs all the radiation that strikes it, and reflects none.

A perfectly reflective surface will have twice as much momentum transferred to it by incident electromagnetic radiation as an illuminated blackbody does, since each wave reverses direction as it reflects off the surface. (This is akin to modeling how gas molecules that collide elastically with a surface collectively exert pressure on it; their **change** in velocity is [negative] two times their initial velocity, so their change in momentum is [negative] two times their initial momentum.) Equation 2 states that the radiation pressure exerted on a perfect reflector is proportional to the intensity of the radiation falling on it: $P = 2I/c$. The factor of two in the proportion corresponds to the assumption that the radiation is reflected: For a blackbody the analogous equation would be $P = I/c$.

Pushing a spaceship is one way that light's pressure can be applied. Scientists have also shown how laser light can be used to **cool** a gas. The collisions of light waves with the molecules of the gas are used to reduce the momenta of these particles, which reduces their thermal velocities. A system of six lasers is able to create a cubical region containing "optical molasses" that traps gas molecules with a temperature as low as 2 μK (two millionths of a Kelvin above absolute zero). This work won a Nobel Prize in 1997 for Steven Chu, Claude Cohen-Tannoudji, and William Phillips.

Farther from the Earth, the effects of sunlight pressure on the microscopic particles boiling off a comet as it passes close to the Sun are spectacular: The particles form a "tail" that streams off the comet away from the Sun like a windblown plume of smoke, no matter what direction the comet is moving.

Derivation We will use Maxwell's impulse equation, shown in Equation 1, to derive the equation relating pressure and intensity, shown in Equation 2. We will first derive it for an object that absorbs all the radiation energy falling on it (a blackbody); then, the pressure is doubled for a perfect reflector.

Variables

pressure of electromagnetic radiation	P
area of object subject to radiation	A
force exerted by radiation	F
momentum of radiation	p
time interval	Δt
total energy of radiation	U
the speed of light	$c = 3.00\times10^8$ m/s
intensity of radiation	I

Strategy

1. Express the radiation pressure on the object as force divided by area, and further express the force as the average rate of change of momentum.
2. Use Maxwell's expression for the impulse of a wave to find the average rate of change of momentum in terms of the energy ΔU transferred by radiation to the absorbing object. Substitute this expression for the rate of change of momentum into the equation for pressure derived in the previous stage.
3. Finally, use the relationship of ΔU to intensity to get the radiation pressure in terms of intensity.

Physics principles and equations

We will use the definitions of pressure and impulse.

$$P = F/A, \quad \Delta p = F\Delta t$$

Maxwell's equation for the momentum transferred to an absorbing object by electromagnetic radiation during the time interval Δt is

$$\Delta p = \Delta U/c$$

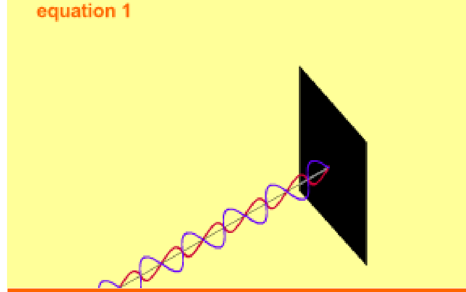
To simplify the derivation, we assume that the rate of energy transfer is constant. Finally, the intensity of electromagnetic radiation is defined as the power it conveys per unit area, and power is the rate of change of energy.

$$I = \frac{P}{A}, \quad P = \frac{\Delta U}{\Delta t}$$

concept 2

Particle theories of light
Light pressure explained by Newton's laws for collisions
· Absorbed light – inelastic collisions
· Reflected light – elastic collisions

equation 1

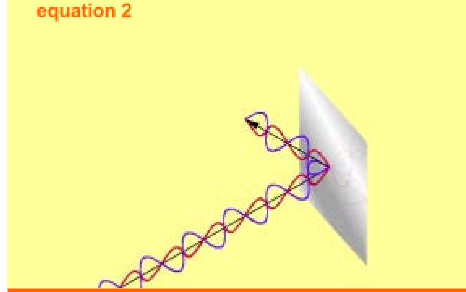


Momentum transferred by radiation absorption (Maxwell)
For a blackbody:

$$\Delta p = \frac{\Delta U}{c}$$

Δp = momentum transferred in time Δt
ΔU = energy transferred in time Δt
c = speed of light

equation 2



Pressure on a perfect reflector

$$P = \frac{2I}{c}$$

P = pressure
I = intensity
c = speed of light

Step-by-step derivation

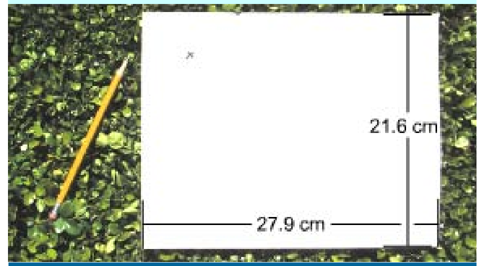
We start with the definition of pressure, and use the definition of impulse to replace the force variable appearing in the equation.

Step	Reason
1. $P = \frac{F}{A}$	definition of pressure
2. $F = \frac{\Delta p}{\Delta t}$	definition of impulse
3. $P = \frac{1}{A}\frac{\Delta p}{\Delta t}$	substitute equation 2 into equation 1

Here, we use Maxell's equation for the momentum transferred to a blackbody absorber by electromagnetic radiation during a time interval Δt. We use this equation to replace the average rate of change of momentum appearing above with an expression that involves energy.

Step	Reason
4. $\Delta p = \frac{\Delta U}{c}$	momentum transferred by radiation
5. $\frac{\Delta p}{\Delta t} = \frac{(\Delta U/\Delta t)}{c}$	divide by Δt
6. $P = \frac{1}{A}\frac{(\Delta U/\Delta t)}{c}$	substitute equation 5 into equation 3

example 1



Sunlight falls with an intensity of 500 W/m^2 on this sheet of paper, which reflects 80% of it. What force does the light exert on the paper?

$P = \frac{1.80\, I}{c}$ (for 80% reflectivity)

$P = \frac{1.80\,(500\ \text{W/m}^2)}{3.00 \times 10^8\ \text{m/s}}$

$P = 3.00 \times 10^{-6}\ \text{N/m}^2$

$F = PA$

$F = (3.00 \times 10^{-6}\ \frac{\text{N}}{\text{m}^2})(0.216\text{m})(0.279\text{m})$

$F = 1.81 \times 10^{-7}\ \text{N}$

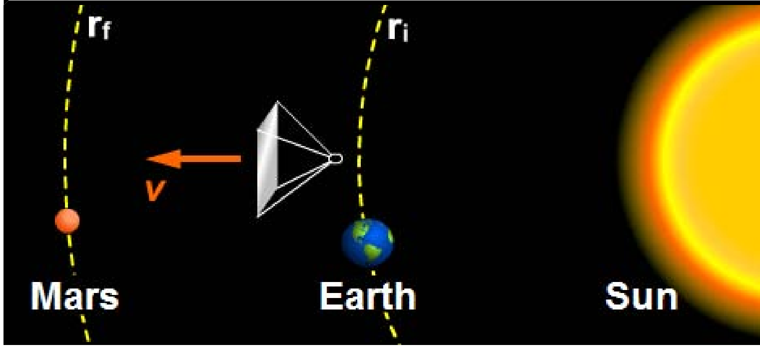
Intensity provides a simpler, more direct way to calculate pressure. We use the relationship of energy and intensity to relate pressure to intensity.

Step	Reason
7. $P = \frac{1}{c}\frac{(\Delta U/\Delta t)}{A}$	rearrange
8. $I = \frac{P}{A} = \frac{(\Delta U/\Delta t)}{A}$	definitions of intensity and power
9. $P = \frac{I}{c}$	substitute equation 8 into equation 7
10. $P = \frac{2I}{c}$	for a perfect reflector

Example 1 asks you to calculate the force exerted by sunlight on a piece of paper. You must first find the light pressure. If the paper were a perfect blackbody, reflecting 0% of the light falling on it, then you would use the formula $P = I/c$. If it were a perfect reflector, reflecting 100% of the incident light, then you would use $P = 2I/c$.

Since the paper in fact reflects 80% of the light falling on it (scientists say it has an *albedo* of 0.8), the appropriate intermediate formula to use is $P = 1.8I/c$. The downward force of the light on the paper is on the order of 10^{-7} N, comparable to the weight of the graphite in the penciled "x" on the upper left corner of the sheet.

You may wonder why, in a discussion of radiation pressure, no mention has been made of the angle at which electromagnetic radiation strikes a surface: Wouldn't oblique radiation exert less pressure than perpendicular radiation? The answer is yes, but this is already accounted for in Equation 2. When radiation striking a surface makes a nonzero angle with its area vector, then the **intensity** of the illumination is itself reduced, being proportional to the cosine of the angle of incidence.



The spacecraft sails away from the Sun, starting at Earth's orbit.

What are the speed and distance traveled by the spacecraft after one hour?

How fast is it going by the time it reaches Mars' orbit?

You have just purchased your brand new Dystis Extranuator space yacht. It carries a vast square sail, 10 kilometers by 10 kilometers, of perfectly reflective material, computer controlled through a harness of InvisiFlex monomolecular cables to have a perpendicular orientation to the incident sunlight at all times. This sleek little puppy, including sails and rigging, has a mass of 9750 kg.

You decide to go for a jaunt heading directly away from the Sun, starting at the distance of the Earth's orbit. How fast will you be sailing after one hour? And how far will you have traveled? For comparison, how fast will you be going when you cross the orbit of Mars?

The power output of the Sun and the orbital distances are given in the table below.

To simplify the problem, we ignore the effects of Earth's gravity, and that of all other bodies. Also, without significant error we may use the initial acceleration of the spacecraft as a "constant" acceleration during the first hour's journey.

First two questions: speed and distance after one hour

Variables

pressure of sunlight on sail	P
intensity of sunlight striking sail	I
the speed of light	$c = 3.00\times10^8$ m/s
power output of the Sun	$P_s = 3.91\times10^{26}$ W
distance from the Sun	r
Earth's distance from the Sun	$r_i = 1.50\times10^{11}$ m
Mars' distance from the Sun	$r_f = 2.28\times10^{11}$ m
force exerted by sunlight on sail	F
area of sail	$A = (1.00\times10^4 \text{ m})^2$
"constant" acceleration of spacecraft	a
mass of spacecraft	$m = 9.75\times10^3$ kg
speed of spacecraft	v
initial speed of spacecraft	$v_i = 0$ m/s
final speed of spacecraft	v_f
elapsed time from start of trip	t
distance traveled from start	Δx

What is the strategy?

1. Use the equations for radiation pressure and intensity to determine the pressure of sunlight on the sail.
2. Use the definition of pressure and Newton's second law to calculate the acceleration of the spacecraft.
3. Then, use two motion equations to determine the final speed and distance traveled by the ship after one hour.

Physics principles and equations

We will need to use the equation for the pressure of electromagnetic radiation on an ideal reflector.

$$P = \frac{2I}{c}$$

The intensity of solar radiation depends on the distance r from the Sun.

$$I = \frac{P_s}{4\pi r^2}$$

Pressure is force per unit area.

$$P = \frac{F}{A}$$

Newton's second law.

$$F = ma$$

We will need two motion equations from the study of kinematics, which we may use because we are assuming that the acceleration of the spacecraft is constant.

$$v_f = v_i + at$$
$$\Delta x = v_i t + \frac{1}{2}at^2$$

Step-by-step solution

First, we find the pressure of sunlight on the sail in terms of the power of the Sun and the distance of the spacecraft from it.

Step	Reason
1. $P = \frac{2I}{c}$	radiation pressure on reflector
2. $I = \frac{P_s}{4\pi r^2}$	radiation intensity
3. $P = \frac{P_s}{2\pi c r^2}$	substitute equation 2 into equation 1 and reduce

We use the definition of pressure, and Newton's second law, to find the acceleration of the spacecraft.

Step	Reason
4. $F = AP$	definition of pressure
5. $F = \frac{AP_s}{2\pi c r^2}$	substitute equation 3 into equation 4
6. $a = \frac{F}{m}$	Newton's second law
7. $a = \frac{AP_s}{2\pi c m r^2}$	substitute equation 5 into equation 6

Now, we calculate the acceleration using the values supplied in the table above. Then we use motion equations to calculate the ship's speed and distance traveled after one hour.

Step	Reason
8. $a = \frac{(1.00 \times 10^4 \text{m})^2 (3.91 \times 10^{26} \text{W})}{2\pi (3.00 \times 10^8 \frac{\text{m}}{\text{s}})(9.75 \times 10^3 \text{kg})(1.50 \times 10^{11} \text{m})^2}$ $a = 9.46 \times 10^{-2} \text{m/s}^2$	evaluate
9. $v_f = v_i + at$	motion equation
10. $v_f = 0 + (9.46 \times 10^{-2} \text{ m/s}^2)(3.60 \times 10^3 \text{ s})$ $v_f = 341 \text{ m/s}$	evaluate
11. $\Delta x = v_i t + \frac{1}{2}at^2$	motion equation
12. $\Delta x = \frac{1}{2}(9.46 \times 10^{-2} \text{m/s}^2)(3.60 \times 10^3 \text{s})^2$ $\Delta x = 6.13 \times 10^5 \text{ m}$	evaluate

Your ship is moving at a good clip at the end of the first hour, about as fast as a jet plane in the Earth's atmosphere, and it has gone a fair distance: 613 km.

Third question: speed at the orbit of Mars. The third question posed above, how fast are you going when your yacht reaches the orbit of

Mars, requires the use of a different motion equation, since the acceleration and the distance traveled are known but the elapsed time is not.

Variables

In the previous steps we calculated the acceleration of the ship at the Earth's orbit, and accurately used this value as its "constant" acceleration for the first hour's journey. Since your acceleration decreases significantly at greater distances from the Sun, but we still wish to use a motion equation that requires constant acceleration, we will use a **mean** acceleration as the "constant" acceleration for the longer journey. We already calculated this value for you at a point halfway between the orbits of the Earth and Mars, using the acceleration formula derived above.

"constant" acceleration of ship	$a = 5.96 \times 10^{-2}$ m/s^2

What is the strategy?

1. Use a motion equation to find the speed of the ship at Mars' orbit.

Physics principles and equations

In this calculation we will use the following motion equation.

$$v_f^2 = v_i^2 + 2a\Delta x$$

Step-by-step solution

We know the acceleration. The distance Δx between the orbits of Earth and Mars can be calculated from the orbital data in the table above. This is the information we need in order to calculate the final speed at Mars' orbit.

Step	Reason
1. $v_f^2 = v_i^2 + 2a\Delta x$	motion equation
2. $\Delta x = 2.28 \times 10^{11}\text{m} - 1.50 \times 10^{11}\text{m}$ $\Delta x = 0.78 \times 10^{11}$ m	difference of orbits
3. $v_f^2 = 2(5.96 \times 10^{-2}\text{m/s}^2)(0.78 \times 10^{11}\text{m})$ $v_f^2 = 0.930 \times 10^{10}\text{m}^2/\text{s}^2$ $v_f = 96{,}400$ m/s	evaluate

As we have observed, the pressure of sunlight on the ship's sail is not really constant. It diminishes as your distance r from the Sun increases. As the pressure of sunlight diminishes, so does the force on the sail, and the acceleration of the spacecraft. In order to take this decrease fully into account, you would need to use calculus.

The correction for nonconstant acceleration is negligible for the modest distance traveled during the first hour of flight. For the trip to Mars' orbit, the correction is larger. In a computation not shown here, we used calculus and found that the actual final velocity of the space yacht is 97,300 m/s. This is greater than the figure yielded by the calculation above, which assumed a constant mean acceleration, but by less than 1%. The time it takes the spacecraft to reach the orbit of Mars (which was not asked for in this problem) is a little more than 16 days.

We also significantly simplified this problem by ignoring the effects of gravity. In fairness to science fiction writers who write about such craft, they typically describe them as being assembled in regions of space safely distant from major sources of gravity. They also often describe how the gravitational forces of planets and stars could be exploited to accelerate ships in their desired directions of travel.

34.15 - Interactive checkpoint: carrier waves

The aircraft carrier USS Enterprise has a topside area of 13,600 m^2, the average albedo of its upper surface is 0.380, and the intensity of the sunlight falling on it is 500 W/m^2.

What is the downward force exerted on the ship by the Sun?

The density of seawater is 1030 kg/m^3. How many cubic centimeters of water must the carrier displace to buoy it up against this force?

Answer:

F = ______ N

V = ______ cm^3

34.16 - How electromagnetic waves travel through matter

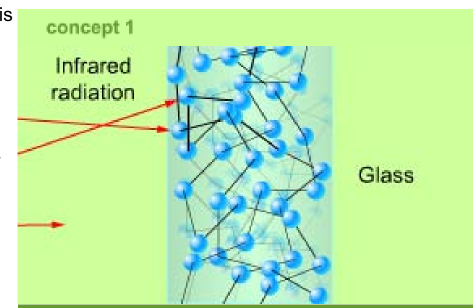
Light and other forms of electromagnetic radiation can travel through a vacuum, and it is often simplest to study them in that setting. However, radiation can also pass through matter: If you look through a glass window, you are viewing light that has passed through the Earth's atmosphere and the glass. Other forms of radiation such as radio waves pass through matter, as well.

This section focuses on how such transmission occurs. It relies on a classical model of electrons and atoms that predates quantum theory. In this model, electrons orbit an atom. They have a resonant frequency that depends on the kind of atom. On a larger scale, atoms themselves and the molecules composed of them also have resonant thermal frequencies at which they can vibrate or rotate.

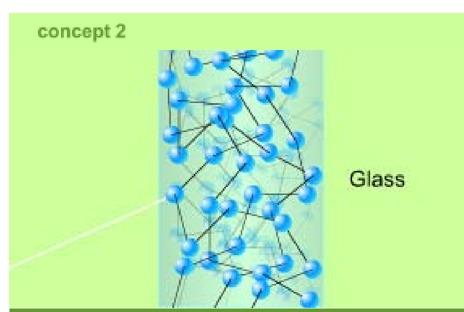
We will use the example of light striking the glass in a window to discuss how substances transmit (or do not transmit) electromagnetic radiation. When an electromagnetic wave encounters a window, it collides with the molecules that make up the glass. If the frequency of the wave is near the resonant thermal frequency of the glass molecules, which is true for infrared radiation, the amplitude of the molecules' vibrations increases. They absorb the energy transported by the wave, and dissipate it throughout the glass by colliding with other molecules and heating up the window. Because it absorbs so much infrared energy, the glass is opaque to radiation of this frequency, preventing its transmission.

Scientists in the 19th century noted a phenomenon in greenhouses caused by the opaqueness of glass to infrared radiation, which they called the *greenhouse effect*. The glass in a greenhouse admits visible light from the Sun, which is then absorbed by the soil and plants inside. They reradiate the solar energy as longer infrared waves, which cannot pass back out through the glass and so help warm up the greenhouse. The same phenomenon occurs on a vaster scale in the atmosphere as gases like methane and carbon dioxide trap solar energy near the Earth's surface.

In contrast to infrared radiation, higher frequency radiation such as visible light does not resonate thermally with atoms or molecules, but may resonate with the electrons of the atoms of a substance. In glass, visible light experiences much less reduction in the amplitude of its waves than infrared radiation does, and most of its energy passes through the glass quite easily. Atoms with resonant electrons that do absorb energy from a light wave quickly pass on that energy by re-emitting it as radiation of the same frequency to other atoms, which in turn pass it on to their neighbors.



Glass opaque to infrared
Radiation is in resonance with molecules
Waves do not travel through



Glass transparent to visible light
Waves absorbed by electrons of atoms
Re-emitted from neighbor to neighbor
Waves pass through but are slowed

This chain of absorptions and re-emissions, called *forward scattering*, follows a path close to the light's original direction of travel. A beam of light that strikes a pane of glass will reach the "last atom" on the far side of the pane in an extremely short time. We see the light after it emerges, and think of glass as transparent.

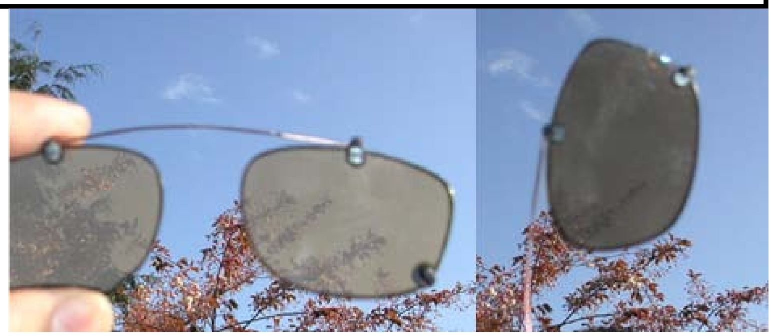
This process does slow the transmission of the wave, which is why light travels slower in glass than it does in air or a vacuum (a fact captured numerically by the *index of refraction* of glass). For instance, light travels through a typical piece of optical glass at about 2/3 of its speed in a vacuum. Of course, 2/3 of the speed of light in a vacuum is still a rather rapid pace….

If atoms of certain substances, such as cobalt, are added to glass, they may absorb certain frequencies of light without re-emitting them. Cobalt glass has a deep blue-violet color, which indicates that all the lower visible frequencies (from red through green) are absorbed and cannot pass through it. Substances which absorb all frequencies of visible light are called *opaque*.

34.17 - Polarization

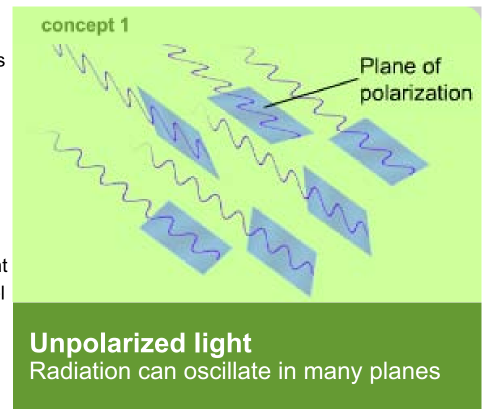
Polarized wave: A transverse wave that oscillates in a single plane.

Polarized radiation: A form of radiation in which the electric field of every wave oscillates in the same plane.



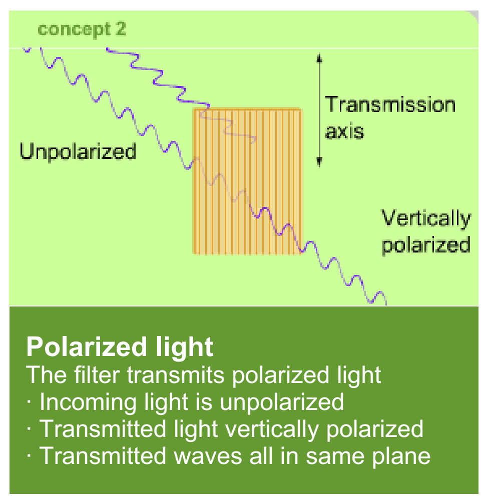
Polaroid sunglasses block some light waves.

Polaroid sunglasses like the ones shown above reduce the amount of light that passes through them. How do they do it? They only let through light waves whose electric fields oscillate in a certain plane. When a source such as the Sun emits light, waves emerge whose electric fields vibrate in every plane parallel to the direction of propagation. This is shown in Concept 1. The electric field of each individual wave does oscillate consistently in its own plane, which is called the wave's plane of polarization, but the radiation as a whole does not have this property.



The electric field of any electromagnetic wave is a vector quantity. A polarizing lens or filter works by only letting through a certain **component** of the electric field of every light wave that strikes it. An ideal polarizing filter can be visualized as a set of narrow parallel slits whose direction is the filter's *transmission axis*.

When a wave passes through the filter, the component of its electric field parallel to the transmission axis is what passes through. The component perpendicular to the axis is absorbed. As a result, waves that oscillate parallel to the slits pass through unhindered, while waves that oscillate perpendicularly to them are completely absorbed.You may want to consider an analogy: a rope passing through a gap in a picket fence. If you shake the rope vertically, the wave you create passes through unhindered. If you shake the rope horizontally, the wave collides with the pickets, transfers energy to them, and does not pass through. In the case of a wave passing through at an oblique angle, the component of its transverse displacement along the "picket axis" passes through while the fence absorbs the other component.



This basic model of how polarization works is shown with an ideal polarizing filter in Concept 2. In the diagram, the transmission axis happens to be **vertical**. A "slit" allows waves oscillating in a vertical plane to pass through. For waves that oscillate in other planes, only the vertical component of the electric field of the wave can pass through. Waves with a horizontal plane of polarization cannot get through the slit at all.

Ordinary light consists of waves whose electric fields are randomly oriented in all lateral directions. This is called *unpolarized* radiation. Two common sources of unpolarized light are the Sun and incandescent light bulbs. If the radiation is created or filtered so that it has only waves oscillating in a single plane, then it is *linearly polarized*. In the illustration in Concept 2, the polarizing filter is exposed to unpolarized radiation, and it transmits linearly polarized radiation.

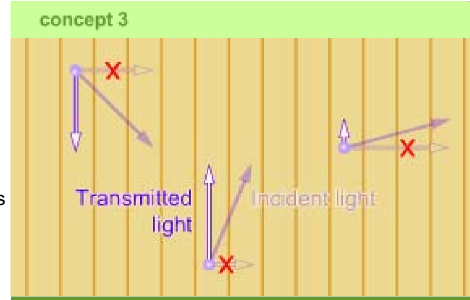
In Concept 3 you see an end-on "close up" of several light waves striking a polarizing filter that has a vertical transmission axis. The light waves are coming toward you. In each case, the vertical component of the electric field of the light is transmitted, and the filter absorbs the horizontal component. The original field is shown as a solid, dimmed vector; its components are hollowed out. The vertical component that passes through is drawn with a bright color, and the horizontal component that is blocked is dimmed and marked with a red ×. In each example illustrated, the electric field amplitude of the transmitted light, and with it the light's intensity, is reduced.

The final illustration, in Concept 4, displays an experiment with two polarizing filters. Unpolarized light is coming toward you from a distant source. It passes first through the upper filter, which allows the passage of light that is polarized at the angle shown by the "slit" lines. This polarized light continues toward you and passes through the lower filter.

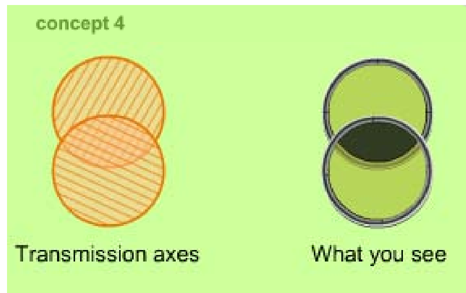
The left-hand part of the illustration shows the orientation of the filters, with parallel lines indicating the transmission axis of each filter. The right-hand side shows you the amount of light that passes through the area of overlap, and how that changes with the angles between the two axes. When the axes are perpendicular, no light at all can pass through.

Radiation also can be *partially polarized*, having a few waves oscillating in all planes, but with most of its waves concentrated in a single plane. This is true of sunlight scattered by the atmosphere. As the photo above shows, the sky in certain directions is partially polarized in a vertical plane so that most of its light can pass through a pair of sunglasses whose transmission axis is vertical. Less light (but still some) passes through the rotated sunglasses. (Polarizing sunglasses are specifically intended to reduce horizontally polarized glare reflected from roadways and water, not skylight.)

Many forms of artificial electromagnetic radiation are polarized. A radio transmitter emits polarized radiation. If the rods of its antenna are vertical, then so is the electric field of every radio wave it creates. In this case, the most efficient receiving antenna is also vertically oriented; a horizontal receiving antenna would absorb radio waves much less efficiently. You may be familiar with this fact if you have ever tried to maneuver a radio antenna wire or a set of television "rabbit ears" to get the best reception. (If you do not know what "rabbit ears" are for television, well, before there was cable television, there was….)



E-field component passes through filter
If transmission axis is vertical, filter …
· Transmits vertical E-field component
· Absorbs horizontal E-field component

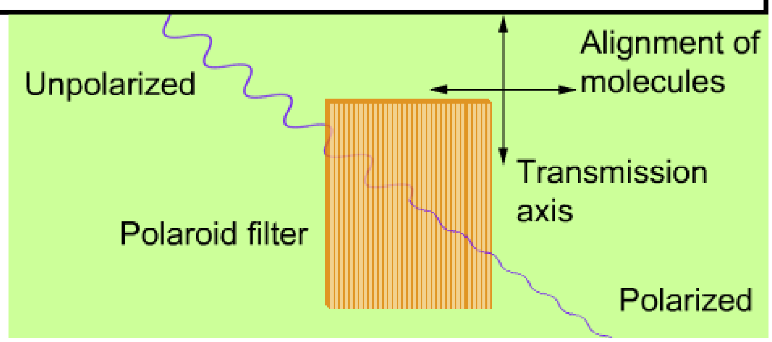


Two Polaroid filters
Light passes through back (upper) filter
Then passes through front (lower) filter
Intensity depends on orientation of filters

34.18 - Polarization and intensity

A polarizing sheet, as found in some sunglasses, reduces the intensity of the light that passes through it. A common form of such material is called *Polaroid*. In this section, we discuss the structure of this material and then review its effect on intensity.

Edwin Land began studying this type of plastic when he was 17 and a student at Harvard University. He dropped out, returned to school, and then left again. In 1937 he and his colleagues formed the Polaroid Corporation, long famous for its instant cameras but well known for its sunglasses, as well.



This filter transmits only the vertical electric field component of waves.

Polaroid material consists of molecules that are long chains of atoms, stretched during manufacture so that they all line up in the same direction. Electrons can move freely along these chains and are able to absorb the energy of electric fields oscillating in the direction of the chain. The result is that electromagnetic waves whose electric fields are parallel to the Polaroid molecules are absorbed, while those with perpendicular fields pass through freely. This means that the transmission axis of the material is perpendicular to the direction of alignment of the molecules, as indicated in the diagram above.

In the previous section, we showed what happens qualitatively when unpolarized light, coming toward you, passes through two polarizing filters. This experiment is shown again in Equation 1. The top filter, through which the light passes first, is called the *polarizer* : The light is polarized after passing through this filter. The bottom filter, through which the light passes subsequently, is called the *analyzer*.

When the transmission axes of the filters are parallel, the polarized light from the polarizer has no trouble passing through the analyzer. As the analyzer rotates, the amount of light able to pass through both filters steadily decreases. When the transmission axes of the filters are at right angles, a configuration referred to as "crossed," no light gets through at all. At this point, all the polarized light coming from the polarizer is perpendicular to the transmission axis of the analyzer and it gets completely absorbed there.

The results of this experiment can be expressed by two equations, which you see in Equation 1. The angle θ is measured between the transmission axis of the analyzer and the axis of the polarizer behind it. The light from the polarizer has an electric field with amplitude E_p. The component of this electric field parallel to the analyzer's transmission axis is $E_p \cos\theta$. This component, which we have designated E_a, passes through the analyzer.

The equation should confirm how you would expect two polarizing materials to function. The polarizer allows electric fields with a specific orientation to pass through. If the angle between the transmission axes of the polarizer and the analyzer is zero (the axes are parallel) then this electric field passes through undiminished. Mathematically, the cosine of zero is one, so the equation confirms this analysis. Conversely, if the axes are perpendicular ($\theta = 90°$) then no wave can pass through both filters, and the amplitude of the electric field of the doubly filtered wave is zero. The cosine of $90°$ is zero, so again the equation confirms what a qualitative understanding of polarization would lead you to believe.

It is also possible to express the **intensity** I_a of the light transmitted by the analyzer in terms of the intensity I_p of the polarized light coming to it from the polarizer. This is the second equation you see in Equation 1. It is named after the French engineer Louis Malus (pronounced "mah-LOO"), who discovered it. The law is a direct consequence of the equation above it, since $I = E^2/2\mu_0 c$.

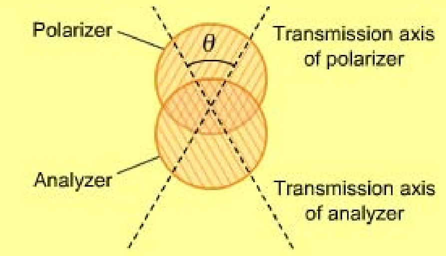
Equation 2 provides a formula for calculating the intensity of unpolarized light as it passes through a **single** polarizing filter. The incoming light has waves whose electric fields, of amplitude E_0, are randomly directed, making various angles θ with the transmission axis of the filter. The resulting transmitted intensity for each individual wave is $I_0 \cos^2 \theta$. The overall transmitted intensity is the average, over all angles θ, of the intensities of these waves. Since the average value of a squared cosine function over all angles is exactly ½, the intensity after the polarizing filter is half the original intensity.

That is, when unpolarized light passes through a filter, the intensity of the transmitted polarized light is one-half the intensity of the incident unpolarized light. For this reason, Polaroid sunglasses, which are designed to completely block horizontally polarized glare, at the same time have the effect of cutting in half the overall intensity of ambient, unpolarized daylight reaching the eyes.

The example problem asks you to obtain a surprising result: More light may pass through three filters than two. Two polarizing filters crossed at $90°$ transmit no light at all. In the problem, we introduce a third, middle filter whose transmission axis is "halfway between" the others, at an angle of $45°$. If this middle filter is inserted **between** the polarizer and the analyzer, the amount of light passing through the analyzer actually increases!

This is because when polarized light from the polarizer encounters the middle filter, the component of each wave that is parallel to the middle filter's transmission axis passes through. So, the light passing through the combination of the first two filters is diminished (but not stopped), and it has a polarization angle of $45°$. This polarized light encounters another "$45°$ obstacle" when it reaches the analyzer, which it passes through diminished again, but not stopped. In this way, polarizing filters behave quite differently than dye or pigment filters, which always have a strictly subtractive effect on the transmission of light passing through them.

equation 1



Polarized light passing through an analyzer

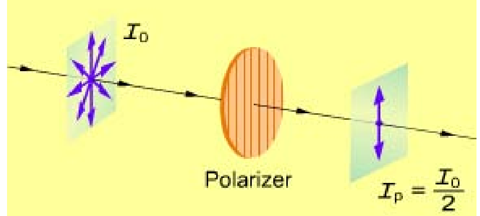
$$E_a = E_p \cos\theta$$

Malus' law:

$$I_a = I_p \cos^2\theta$$

E_p = **E** field amplitude after polarizer
E_a = **E** field amplitude after analyzer
θ = angle between polarizer, analyzer
I_p = intensity transmitted by polarizer
I_a = intensity transmitted by analyzer

equation 2

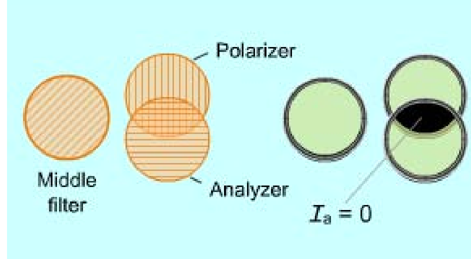


Unpolarized light passing through a polarizer

$$I_p = I_0 / 2$$

I_0 = intensity of unpolarized light

example 1



The "middle" filter is inclined at 45°. If it is placed between the "crossed" polarizer and analyzer filters, what is the new value of I_a?

polarizer-to-middle: θ is 45°

$I_m = I_p \cos^2(45°)$

$I_m = I_p / 2$

middle-to-analyzer: θ is 45°

$I_a = I_m \cos^2(45°)$

$I_a = (I_p/2) / 2 = I_p / 4$

In terms of unpolarized light I_0,

$I_a = (I_0/2) / 4 = I_0 / 8$

34.19 - Sample problem: gazing at the Sun

Light that delivers more than 500 μW of power to the human eye can severely damage the retina. Is it safe to stare at the Sun from Mars as long as you are wearing Polaroid sunglasses?

Mars' atmosphere is so thin that you can ignore its effect on the intensity of sunlight. Assume that the pupil admitting light into the human eye is a circle with a diameter of 2.00 mm.

Variables

intensity of sunlight at Mars' orbit	I_0
radius of Mars' orbit	$r = 2.28\times10^{11}$ m
power emitted by the Sun as light	$P_s = 3.91\times10^{26}$ W
intensity transmitted by polarizer	I_p
power delivered into eye	P
diameter of pupil	$d = 2.00\times10^{-3}$ m
area of pupil	A

What is the strategy?

1. Use the power P_s of the Sun to find the intensity I_0 of its visible electromagnetic radiation (white light) at the distance of Mars' orbit.
2. Calculate the intensity I_p of the Sun's visible light that passes through a polarizing filter on Mars.
3. Calculate the power P of the polarized sunlight that passes through an area the size of the pupil of a human eye on Mars.

Physics principles and equations

For an omnidirectional source of electromagnetic radiation,

$$I = \frac{P}{4\pi r^2}$$

For unpolarized light of intensity I_0 that strikes a polarizing filter, the transmitted intensity I_p is

$$I_p = \frac{I_0}{2}$$

The intensity of electromagnetic radiation is measured in W/m^2. The power P absorbed by an area A illuminated by light of intensity I equals the product IA.

Step-by-step solution

In the first steps we find an expression for the intensity I_p of the sunlight passing through a polarizing filter at the distance of Mars from the Sun. We also state an expression for the area of the pupil of the human eye.

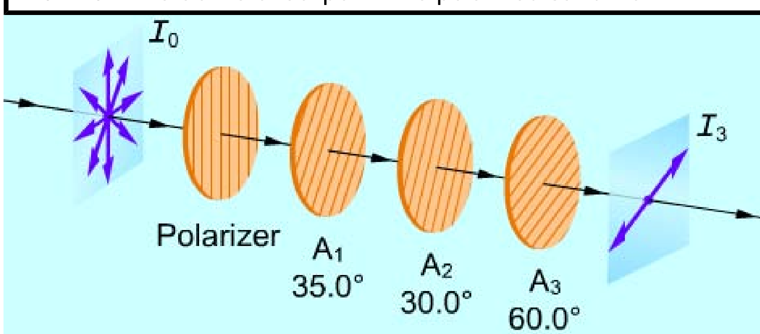
Step	Reason
1. $I_0 = \dfrac{P_s}{4\pi r^2}$	radiation intensity and distance
2. $I_p = \dfrac{I_0}{2}$	intensity transmitted by polarizer
3. $I_p = \dfrac{P_s}{8\pi r^2}$	substitute equation 1 into equation 2
4. $A = \pi\left(\dfrac{d}{2}\right)^2 = \dfrac{\pi d^2}{4}$	area of circle

We use the expressions derived above to find an expression for the power of the polarized sunlight entering the human eye on Mars, and in the last step we evaluate this expression.

Step	Reason
5. $P = I_p A$	power is intensity times area
6. $P = \left(\dfrac{P_s}{8\pi r^2}\right)\left(\dfrac{\pi d^2}{4}\right)$	substitute equations 3 and 4 into equation 5
7. $P = \dfrac{P_s d^2}{32 r^2}$	simplify
8. $P = \dfrac{(3.91\times10^{26}\text{ W})(2.00\times10^{-3}\text{ m})^2}{32\,(2.28\times10^{11}\text{ m})^2}$ $P = 9.40\times10^{-4}\text{ W} = 940\ \mu\text{W}$	evaluate

As you can see, the power of the electromagnetic radiation entering the observer's eyes substantially exceeds the safety limit of 500 μW. Even with Polaroids, the Sun is too bright to look at from the distance of Mars, although the Red Planet is half again as far from the Sun as the Earth is. As for the dubiously authentic experiment depicted above: Do not try this on your home planet!

34.20 - Interactive checkpoint: the polarized sandwich



Unpolarized light having intensity I_0 strikes a polarizer filter, then passes sequentially through three analyzer filters A_1, A_2, and A_3. These filters are rotated clockwise through the angles shown with respect to the orientation of the polarizer. What is the intensity I_3 of the light transmitted by A_3, in terms of the original intensity I_0?

Answer:

$I_3 =$ ______ I_0

Scattering: Absorption and re-emission of light by electrons, resulting in dispersion and some polarization.

Scattered sunlight gives the sky its blue glow.

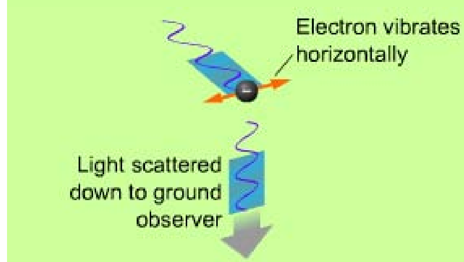
The answer to a classic question − Why is the sky blue? − rests in a phenomenon called scattering. In this section, we give a classical (as opposed to quantum mechanical) explanation of how scattering occurs.

When light from the Sun strikes the electrons of various atoms in the Earth's atmosphere, the electrons can absorb the light's energy, oscillating and increasing their own energy. The electrons in turn re-emit this energy as light of the same wavelength. In effect, the oscillating electrons act like tiny antennas, emitting electromagnetic radiation in the frequency range of light.

An electron oscillates in a direction parallel to the electric field of the wave that energizes it, as shown in Concept 1. The electron then emits light polarized in a plane parallel to its vibration. We show a particular polarized wave that is re-emitted downward toward the ground, since we are concerned with what an observer on the surface of the Earth sees. Other light is scattered in other directions, including light scattered upward and light scattered forward in its original direction of travel.

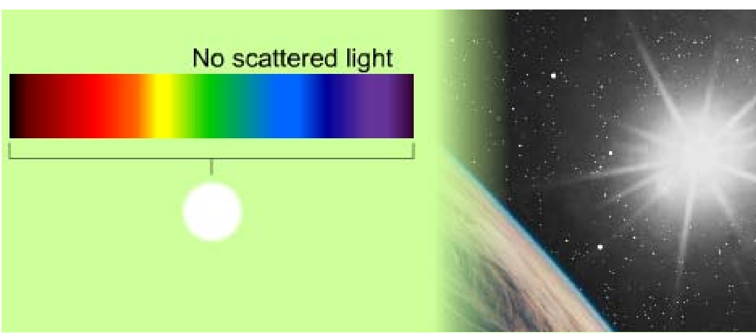
Scattering explains why we see the sky: Light passing through the atmosphere is redirected due to scattering toward the surface of the Earth. In contrast, for an astronaut observer in the vacuum of space, sunlight is not scattered at all so there is no sky glow: Except for the stars, the sky appears black. To the astronaut, the disk of the Sun, a combination of all colors, looks white. We illustrate this below: The full spectrum combines to form white light.

concept 1



Scattering

Light from Sun hits electron
· Electron oscillates
· Re-emits polarized wave to ground
· Shortest wavelengths scattered most



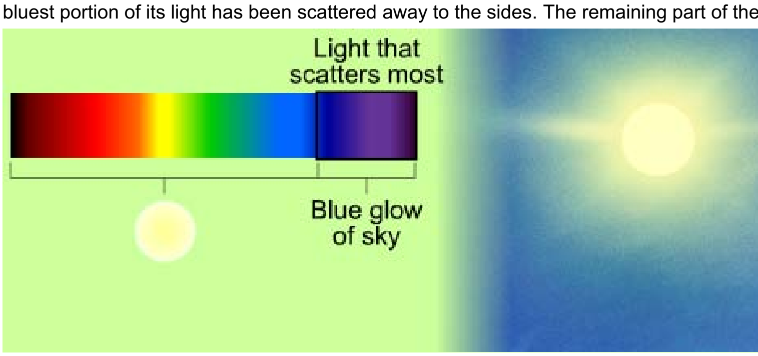
concept 2

View from space

No scattering: sky is black
Sun appears white

The question remains, why is our sky blue rather than some other color? Light at the blue end of the visible spectrum, which has the shortest wavelength, is 10 times more resonant with the electrons of atmospheric atoms than red light. This means blue light is scattered more than red, so that more of it is redirected toward the ground.

Scattering also explains why we see the Sun as yellow rather than white. When you look up at the disk of the Sun from the Earth's surface, the bluest portion of its light has been scattered away to the sides. The remaining part of the Sun's direct light appears somewhat yellowish.

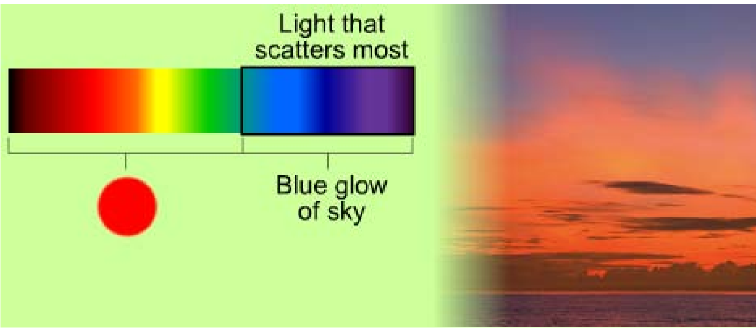


concept 3

Why the sky is blue (and the Sun is yellow)

Shortest waves are scattered: sky is blue
With blue scattered, Sun appears yellowish white

You may also have noted how the Sun appears to change color when it sets. As the Sun's disk descends toward the horizon, its light must pass through a greater and greater thickness of atmosphere in order to reach you. Since a certain amount of sunlight is scattered aside for each kilometer of atmosphere it passes through, its position at sunset causes it to lose large amounts of light at the blue end and even toward the middle of the visible light spectrum. At sunset, practically all the shorter wavelengths of light have been scattered out of it, leaving only light at the red end of the spectrum to be viewed by you. The “missing” blue light is not really missing. People to the west of you perceive it as the daytime sky.



concept 4

Sunset
Longer wavelengths also scattered
Sun appears red

Scattering also explains why skylight is partially polarized. When the Sun is low in the sky, as depicted in Concept 1, horizontally polarized light that gets scattered down from the overhead sky is polarized in the plane of the downgoing wave you see in the illustration. Incoming sunlight that is polarized in other planes also gets scattered, but not straight down towards the ground.

You can experiment with skylight polarization yourself if you have a pair of polarizing sunglasses. In the early morning or late afternoon hold your glasses against the northern or southern sky at arm’s length. Turn one of the lenses slowly, recalling that its transmission axis is vertical when the sunglasses are worn normally. You will find that the skylight is partially polarized in a plane perpendicular to the direction to the Sun.

34.22 - Optically active substances

Optically active substance: One that rotates the plane of polarization of light passing through it.

Substances that change the direction of polarization of light passing through them are called *optically active*. In this section we will discuss applications of optical activity in solids and liquids: *stress analysis* and *polarimetry*.

Certain transparent substances, such as the plastics Lucite® and Plexiglas®, are not normally optically active, but they become so when they are stressed. The greater the resulting strain in the plastic, the more it rotates polarized light; more precisely, the more it rotates the plane of polarization of the light. These kinds of materials possess an *optic axis*, which behaves quite differently from a transmission axis. The **speed** of polarized light passing through them depends on the orientation of the light’s plane of polarization with respect to the optic axis, and the plane of polarization is gradually rotated toward or away from the optic axis as the light passes through.

When a stressed plastic is placed between a polarizer and a crossed (perpendicular) analyzer and illuminated, the transmitted light will be zero where there is no strain (no rotation occurs) and brightest where there is the most strain (the most rotation occurs). The longer the wavelength of light, the less its plane of polarization rotates for a given strain, and the more strain it takes to rotate it through an angle that will allow it to pass through the analyzer. An example of this is shown in Concept 1: In the plastic model, the red regions are showing the most strain, while the violet and blue regions are showing the least strain.

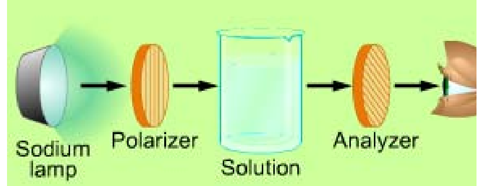
The hands in the photograph, which are illuminated from behind, appear black. The polarized background would normally appear black too; in this photograph it has been brightened by the addition of extra cross-polarized light that is able to pass through the analyzer.

Engineers often use the optical activity of transparent solids in a procedure called optical stress analysis. Before building a bridge, for example, they may build a plastic model of the bridge and subject it to various loads. By using polarized light to observe how the strains imposed by a load distribute themselves over the structure, they can discover which parts of it are the most vulnerable to stress.

concept 1

Optically active solids
Optical activity proportional to strain
· Red regions: most strain
· Violet regions: least strain

concept 2



Optically active solutions
Polarimeter measures rotation of polarized light
Depends on path length, concentration
Proportionality constant: specific rotation

Many substances, when dissolved in a liquid such as water or alcohol, form a solution that rotates polarized light. For example, if a beam of polarized light passes through a solution of sugar water, its plane of polarization may be rotated either clockwise or counterclockwise. A solution of grape sugar, or *dextrose*, rotates polarized light to the right (clockwise). In fact, its name comes from the Latin word for right. The more healthful fruit sugar fructose, sometimes called *levulose* from the Latin word for left, rotates polarized light in the opposite direction.

Why does this rotation occur? When a light wave is absorbed and re-emitted by an electron in a dissolved asymmetrical molecule such as dextrose, the forward scattering imparts a slight twist to its plane of polarization. Optically active molecules in a solution contain complex asymmetric electric fields that can change the plane of polarization of a wave.

The greater the number of molecules a light wave interacts with in an optically active substance, the more its plane of polarization rotates. For this reason, the overall rotation of polarized light passing through a solution depends on the length of the light path through the solution and on the concentration of the dissolved substance. The rotation is also proportional to a constant α_0 called the *specific rotation* of the substance, which reflects the rotating power of its molecules. These relationships are summarized to the right in the *polarimeter equation*. Note that (and this is unusual for a physics equation) the rotation angle α is measured in degrees rather than radians, and the clockwise direction is considered positive.

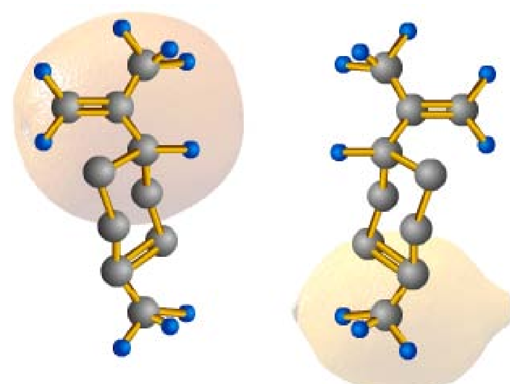
The *polarimeter* is a device that can be used to measure the net rotation of polarized light passing through an optically active solution. An experimenter directs polarized light through a container of the solution to be analyzed. The analyzer, which starts out parallel to the polarizer, does not transmit all the light from the polarizer because the light's plane of polarization has been rotated by the solution. The experimenter turns the analyzer to one side or the other until the transmitted light has maximum brightness. Then she knows that the analyzer's transmission axis matches the rotated polarized light, and she can measure the angle α through which the analyzer has turned.

The polarimeter equation gives an expression for the angle α of the analyzer at which the transmitted light will be the brightest. If the polarized light encounters more molecules of the optically active substance, either because the solution is more concentrated or because the immersed light path is longer, the rotation will be greater. Since the amount of rotation also depends on the wavelength of the light, the specific rotations α_0 given in tables for particular dissolved substances are based on a polarimeter employing the 589 nm light that is emitted by a sodium vapor lamp.

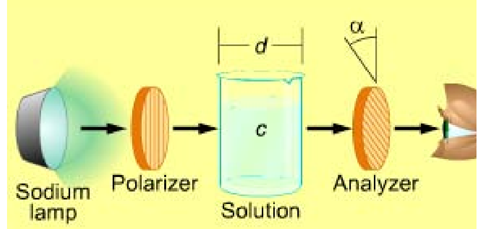
Dextrose and fructose molecules are chemically identical (they have the same atoms arranged in the same pattern) but they are mirror images of each other. Because of this they rotate polarized light by the same amount in opposite directions. Organic molecules such as *carvone* may exist in two mirror image forms; you smell carvone as caraway or spearmint, depending on which way the molecule twists. The scents are different because the smell receptors in the nose react differently to the mirror image forms.

Using a polarimeter is one way to distinguish between the two forms of mirror image compounds. Also, if the specific rotation of a particular substance is known, the device can be used together with the polarimeter equation to determine the concentration of the substance in a solution. You are asked to perform such an analysis in the example problem to the right.

The diagrams below show the mirror image molecular forms of the citrus oil *limonene*, which is the essence of either orange or lemon, depending on the orientation of its molecules! (The gray spheres represent carbon atoms, and the blue spheres are hydrogen atoms.)



equation 1

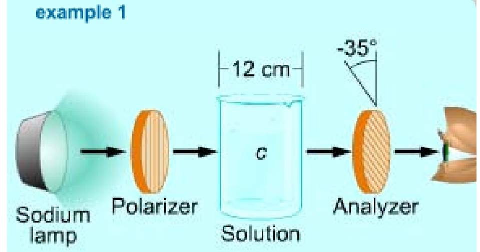


Polarimeter equation

$$\alpha = d c \alpha_0 / 100$$

α = rotation of light (° clockwise)
d = length of immersed light path (m)
c = concentration of substance (kg/m^3)
α_0 = specific rotation of substance
Units of α_0: °m^2/kg

example 1



Carvone's specific rotation is +62.5 (caraway) or –62.5 (spearmint). What is the concentration of the carvone in this beaker?

$\alpha = d c \alpha_0 / 100$

$$c = \frac{100\alpha}{d\alpha_0}$$

Counterclockwise rotation means spearmint, so we use $\alpha_0 = -62.5$:

$$c = \frac{(100)(-35°)}{(0.12\text{ m})(-62.5\ °\text{m}^2/\text{kg})}$$

$c = 470$ kg/m^3

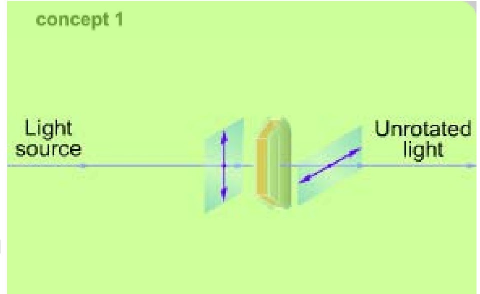
34.23 - Physics at work: liquid crystal displays (LCDs)

The liquid crystal display (LCD) in the watch face above demonstrates *variable optical activity* at work. LCDs are found in many common devices, including calculators, cellular telephones and clocks. There are two types of LCDs: backlit and reflective. The one shown above is a reflective LCD, but we will explore both types in this section. Backlit LCDs generate light behind their displays; reflective LCDs like the one above utilize ambient light.

LCDs rely on polarization. The characters of a digital watch display consist of "digit" segments: regions that can be made dark. These segments are filled with a substance called *liquid crystal* that is optically active

This digital watch displays the time with a reflective LCD.

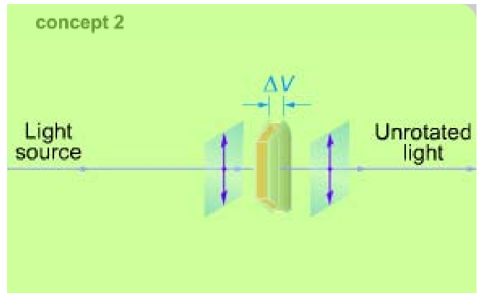
in its natural state but becomes inactive when a potential difference is applied across it. You see this phenomenon illustrated for a backlit LCD in Concepts 1 and 2. In Concept 1, the power is off (there is no potential difference across the crystal), so the crystal is optically active and rotates the plane of polarization of the light. The thickness of the crystal is designed to rotate the light by $90°$. In Concept 2 a potential difference is applied, the crystal becomes inactive, and there is no rotation of the light's plane of polarization.



Power OFF
Liquid crystal is optically ACTIVE
Rotates polarized light 90°

The liquid crystal is sandwiched between two polarizing filters with perpendicular tranmission axes, as shown in Concept 3. In the backlit LCD shown there, a light behind the display shines through the left-hand filter. Following our usual practice we will call this filter that is closer to the light source the polarizer, and the one farther away the analyzer. The polarizer allows light with a particular plane of polarization to pass through. When the liquid crystal is "on" (optically inactive), it does not rotate the polarized light, and the analyzer prevents any of the light from passing through, creating a dark area.

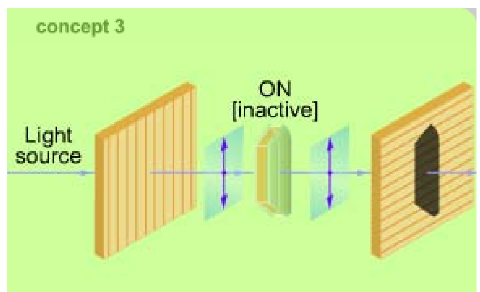
When **no** potential difference is applied across the liquid crystal, it is "off" and reverts to its normal optically active state. The perpendicular orientation of the analyzer now allows the polarized light to pass through. Instead of being black, the material of the segment looks the same as the adjoining material. In order to make the transmitted light difficult to see, the display's background is colored to resemble it. This is shown in Concept 4.



Power ON
Liquid crystal is optically INACTIVE
Does not rotate polarized light

If you examine the watch in the illustration above or one on your wrist, or a friend's wrist, you will see that the digits it displays are pieced together from segments that appear black when they are "turned on." When they are off, these segments are invisible to the casual glance, but they can still be made out as faint shadows if you look very carefully.

We have been describing a typical backlit LCD, which can be read in the dark. The backlighting consumes more power than any other part of the display, since the amount of power required to turn on each digit segment is negligible. Backlit LCDs are typically found in automobile dashboards, where their power usage is not a particular concern and where readability at night is important.

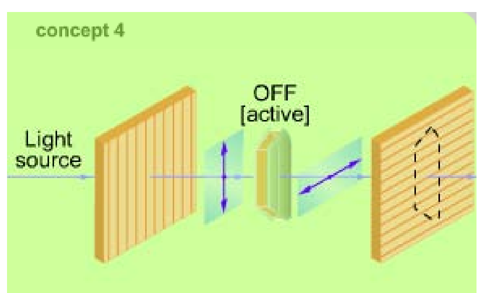


Backlit LCD – ON
Analyzer absorbs unrotated light
· Segment appears black

A reflective LCD, such as the one in the watch above, consumes less power. We show how a reflective LCD works in Concept 5; it is a bit more complicated than a backlit LCD. A mirror that can reflect ambient light coming from in front of the LCD replaces the backlight behind the polarizer. As with the backlit display, the transmission axes of the polarizer and analyzer are at right angles to each other.

Concept 5 shows what happens when a liquid crystal digit segment in this type of LCD is turned **off**. In this state, light passes through each of the analyzer, the liquid crystal segment, and the polarizer twice. Unpolarized light enters the display from the right and becomes horizontally polarized as it passes through the right-hand filter (the analyzer). It rotates $90°$ to a vertical orientation as it passes through the optically active liquid crystal component. In this orientation it can pass through the polarizer, reflect off the mirror, and come back out through the polarizer without hindrance. (A keen-eyed observer will note that the angular symmetry of reflection is apparently being violated where the light strikes the mirror: We drew the incident ray at a different angle in order to fit all the details into the diagram.)

After passing through the polarizer, the light retraces its path through the liquid crystal, again rotating and again becoming horizontally polarized. It passes back out through the analyzer without absorption, creating a region that blends into the background color of the display.

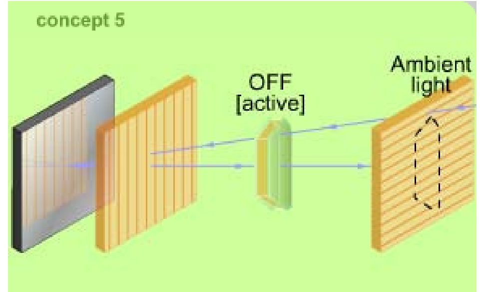


Backlit LCD - OFF
Analyzer admits rotated light
· Segment blends into background

When the segment is turned on, it becomes optically inactive; optically, it can be treated as if it were no longer there. Light entering from the right gets horizontally polarized by the analyzer, and propagates to the left until it strikes the polarizer, where it is absorbed. This light never even reaches the mirror, and the segment appears dark. You see this happening in the watch above, where the dark segments spell out the time of day.

More elaborate LCDs of both types can be manufactured with "segments" having any shape, not just the parts of digits. For example, the battery and signal strength icons on a cellular telephone display, or the letters "H" and "M" on the watch above, use specially shaped segments. Some flat panel display screens use an array of thousands or even millions of tiny LCD dot-segments to produce virtually any image. With the appropriate refinements, color images can be produced.

As with polarized light from the sky, you can use a pair of Polaroid sunglasses to experiment with the polarized light emitted by a liquid crystal display. For example, you will find that a watch, viewed through the glasses, looks quite different depending on its orientation. At some angles the display looks more or less normal, while at others it becomes completely unreadable.



Reflective LCD
When power OFF, segment is not seen
When power ON, segment is black

34.24 - Gotchas

A light wave is a transverse wave. Yes. Both of its components, an electric and a magnetic field, oscillate perpendicularly to its direction of travel.

Radio signals and light waves are fundamentally different. Both are forms of electromagnetic radiation, so we lean toward "no" in response to this statement. The wavelength and frequency of radio transmissions and light are significantly different, and humans can see light, but not radio waves, so one could say "yes". However, both are electromagnetic waves, and both move at the speed of light.

Intensity is the same thing as average energy density. No, intensity represents the average rate of **power** transmission of an electromagnetic wave per unit **area**, perpendicular to the direction of propagation of the wave. Average energy density represents the average amount of **energy** contained in an electromagnetic wave per unit **volume**.

34.25 - Summary

An electromagnetic wave is a traveling wave consisting of mutually perpendicular electric and magnetic fields that oscillate transversely to the direction of propagation. Electromagnetic radiation moves at the "speed of light," or c, which is 299,792,458 m/s in a vacuum. The value 3.00×10^8 m/s is often used.

Every electromagnetic wave has a characteristic frequency and wavelength. The electromagnetic spectrum is an ordering of electromagnetic radiation in accordance with these two properties and extends far beyond the tiny gamut called visible light that we can detect with our eyes. Some other kinds of electromagnetic radiation are radio waves (AM and FM), television signals, microwaves, infrared light, ultraviolet light, x-rays, and gamma rays.

Oscillating electric and magnetic fields propagate as a wave indefinitely, as the changing magnetic field induces a changing electric field (Faraday's law), and the changing electric field in turn induces a changing magnetic field (Maxwell's law of induction). An electromagnetic wave does not require a medium to propagate.

Electromagnetic plane waves are the form of electromagnetic radiation that is the simplest to analyze. With these waves, wave fronts advance through space in a series of parallel planes; they are approximated by radiation from a very distant point source. The electric and magnetic fields that constitute plane waves can be mathematically described by field equations that are sinusoidal functions of position and time. At a fixed point in space, both fields oscillate sinusoidally in time. At a fixed instant in time, both fields can be depicted as mutually perpendicular sinusoidal wave trains.

Maxwell's equations can be used to derive a pair of general wave equations. These differential equations must be satisfied by any function of position and time that describes an electromagnetic wave. The sinusoidal field functions for a plane wave satisfy these general wave equations. The wave equations imply that the electric and magnetic field strengths of an electromagnetic wave are proportional at all points in space and at all times, with the constant of proportionality being c, the speed of light.

Maxwell proved that the speed of an electromagnetic wave equals the reciprocal of the square root of the product of two fundamental physical constants, the electric permittivity of free space ε_0, and the magnetic permeability of free space μ_0. This value turned out to be equal to the empirically well-measured speed of light, providing strong evidence that light is a form of electromagnetic radiation.

The intensity of electromagnetic radiation equals the average power transported by the radiation per unit area, measured perpendicularly to the direction of propagation.

Equations

Proportionality of fields

$$\frac{E}{B} = c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

Poynting vector

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}$$

Intensity of electromagnetic radiation

$$I = S_{\text{avg}} = \frac{E_{\text{max}}^2}{2\mu_0 c} = \frac{E_{\text{rms}}^2}{\mu_0 c}$$

$$I = \frac{P}{4\pi r^2}$$

Energy density

$$u_{\text{E}} = \frac{\varepsilon_0 E^2}{2}, \quad u_{\text{B}} = \frac{B^2}{2\mu_0}$$

$$u_{\text{E}} = u_{\text{B}}$$

$$u = u_{\text{E}} + u_{\text{B}} = 2u_{\text{E}} = 2u_{\text{B}}$$

$$u_{\text{avg}} = \frac{\varepsilon_0 E_{\text{max}}^2}{2}$$

Momentum transferred by radiation absorption

$$\Delta p = \frac{\Delta U}{c} \quad \text{for a blackbody}$$

Intensity is measured in W/m^2. For visible light, it corresponds roughly to what humans perceive as brightness. The energy density of a wave is the amount of energy in the wave per unit volume. Energy density is measured in J/m^3. Both intensity and energy density are proportional to the square of the electric field amplitude E_{max} of the wave.

Electromagnetic radiation emitted isotropically (equally in all directions) by a point source takes the form of spherical waves rather than plane waves. The intensity of these waves is proportional to the power of the source and inversely proportional to the square of their distance from the source.

Electromagnetic waves exert radiation pressure on any surface they illuminate, proportional to the intensity of the radiation. The pressure on a perfectly reflective surface is twice that on a perfectly absorptive surface.

Linearly polarized light consists of light waves whose electric fields all oscillate in the same plane. This type of light can be created by several methods, such as by passing unpolarized light, whose electric fields oscillate in many directions, through a polarizing filter. The direction of polarization that results is called the transmission axis of the filter. As a randomly polarized wave passes through such a filter, only the component of its oscillating electric field that is aligned with the transmission axis passes through.

Polaroid sunglasses take advantage of the fact that a polarizing filter can decrease the intensity of the light that passes through it, and the transmission axes of their lenses are oriented vertically to block the horizontally polarized glare reflected from the surfaces of roadways and bodies of water.

Pressure on a perfect reflector

$$P = \frac{2I}{c}$$

Polarized light passing through an analyzer

$$E_a = E_p \cos\theta$$

Malus' law

$$I_a = I_p \cos^2\theta$$

Unpolarized light passing through a polarizer

$$I_p = I_0 / 2$$

Polarimeter equation

$$\alpha = dc\alpha_0 / 100$$

The scattering of light passing through a transparent substance is the absorption and re-emission of light waves of characteristic frequencies by atoms in the substance. Scattering in the atmosphere is responsible for the blue glow of the sky, the yellowish hue of the Sun, and the red color of sunsets.

Optically active substances rotate the plane of polarization of polarized light that passes through them.

Optical activity forms the technological basis of the liquid crystal displays (LCDs) used in many consumer electronic devices.

Chapter Assumptions

Unless stated otherwise, electromagnetic waves are assumed to be propagating as plane waves through a vacuum.

When converting light-years to meters, assume that a year has 365 days.

The speed of light in a vacuum is 3.00×10^8 m/s.

The Sun radiates energy at the rate of 3.91×10^{26} W.

Conceptual Problems

C.1 In a phenomenon called **diffraction**, waves are able to bend around obstacles. For example, ocean waves can bend around a piling so that it does not cast a down-wave "shadow" of calm water. The shorter the wavelength of a wave, the less able it is to diffract around obstacles. You have experienced this yourself: You cannot see the people in a room down the hall, but you can hear them talking, corresponding to the fact that a typical wavelength of light is about ten million times shorter than that of a typical sound. If you live near the big city, but behind a rather large mountain, which of the urban radio broadcasts are you likely to receive the best: AM or FM? Explain.

○ AM ○ FM

C.2 Half of all the electromagnetic radiation reaching the Earth from the Sun lies in the visible spectrum that can be perceived by the human eye: ranging from extremes of 390 nm to 780 nm. This is a tiny portion of the whole electromagnetic spectrum, and it can't be just a coincidence. Why is so much of the Sun's radiation visible?

C.3 Electromagnetic radiation is called **polarized** if every individual electromagnetic wave in the radiation has its electric field oscillating in the same plane. Is the radio-wavelength radiation emitted from a dipole antenna polarized? Explain.

○ Yes ○ No

C.4 Clouds scatter light just as the atmosphere does. Since clouds consist mainly of microscopic water droplets of many sizes, they scatter all the wavelengths of visible light equally. The smallest droplets scatter the shortest wavelengths at the violet end of the spectrum, and larger droplets scatter light of longer wavelengths. Since all wavelengths are equally scattered by a cloud, it appears white. So does a jet's condensation trail, or fog, or steam, or mist on a window.

Well, not quite white. In fact if you look at a real (cumulus) cloud, you will see that parts of it appear bright white, while other parts appear gray, and at the bottom of a thick cloud, almost black. What accounts for these variations?

C.5 It is early morning, and the Sun is rising (where else?) in the east. (a) What is the direction of polarization of the skylight directly overhead? (b) What is the direction of polarization of the skylight as you look toward the northern horizon?

(a) i. East-west
  ii. North-south
(b) i. Vertical
  ii. Horizontal

C.6 Now it is noon, and the Sun is directly overhead. (a) What is the direction of polarization of the skylight coming from the eastern horizon? (b) What is the direction of polarization of the skylight as you look toward the northern horizon?

(a) i. Vertical
  ii. Horizontal
(b) i. Vertical
  ii. Horizontal

C.7 Sunlight which reflects off smooth surfaces like those of roadways or standing water becomes horizontally polarized in the process. Polarizing sunglasses are designed to completely block this reflected glare, while at the same time they reduce the intensity of unpolarized ambient light by one-half. (a) If a sunglasses lens is to block reflected glare, what should the orientation of its transmission axis be? (b) LCD displays, such as flat panel computer display screens, and certainly automobile instruments and digital writswatch faces, should be readable by people wearing polarized sunglasses. In order for them to be readable, what should the transmission-axis orientation of their front "analyzer" filters be? (Note: The conceptual diagrams of LCDs in the text may or may not have been drawn to reflect this requirement.)

(a) ○ Vertical ○ Horizontal
(b) ○ Vertical ○ Horizontal

Section Problems

Section 1 - The electromagnetic spectrum

1.1 The speed of light in a vacuum is exactly 299,792,458 m/s. This speed is sometimes used to provide a convenient yardstick for large astronomical distances. (a) A light-second is the distance light travels in one second. If the Moon is 3.84×10^8 m from the Earth (center to center), how many light-seconds away is it? (b) A light-minute is the distance light travels in one minute. If the Earth is 1.50×10^{11} m from the Sun, how many light-minutes away is it? (Your answer also represents the number of minutes it takes for light to reach the Earth from the Sun.) (c) A light-year is the distance light travels in one year. The Sun is 28,000 light-years from the center of the Milky Way galaxy. What is this distance in meters?

(a) ________ light-seconds
(b) ________ light-minutes
(c) ________ m

1.2 Nowadays many highway tunnels are equipped with AM "repeaters" so that motorists can listen to their AM radios while inside. Otherwise, AM radio cannot be received in the tunnel. FM repeaters are not usually necessary, since FM reception in tunnels is often acceptable without any special assistance. A typical tunnel is 5 meters high. (a) What is the wavelength of the electromagnetic radiation from an FM station broadcasting at 94.1 MHz, which penetrates easily into the tunnel? (b) What is the wavelength of the electromagnetic radiation from an AM station broadcasting at 940 kHz, which ordinarily cannot penetrate into the tunnel?

(a) ________ m
(b) ________ m

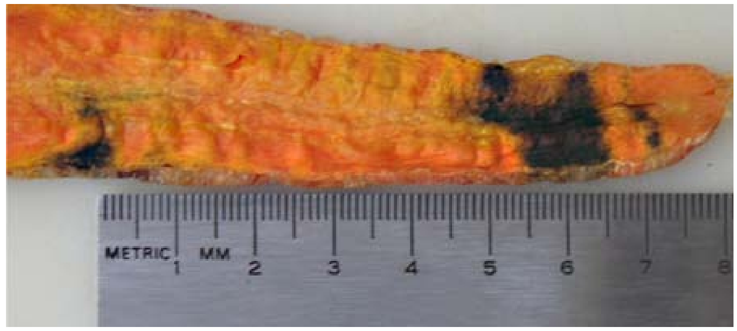
1.3 A certain radio station broadcasts electromagnetic radiation with a wavelength of one quarter mile (0.25 mi). (a) What is the frequency of the broadcast? (b) Is the station AM or FM?

(a) ________ Hz
(b) ○ AM ○ FM

1.4 The color to which the human eye is most sensitive is a yellowish green with a wavelength of approximately 550 nm. (For this "scientific" reason there was a brief vogue in the 1980s for making fire trucks this color, a fad that faltered when departments found that firefighters took little pride in non-red fire trucks and consequently maintained them poorly. In fact, the photographer of the accompanying picture was told that the truck will be painted red "within a month − red is traditional".) (a) What is the frequency of waves of this color? (b) How many wavelengths of this yellowish green light are there in the thickness of a human hair that is 110 μm in diameter? (c) How many wavelengths of the light are there in the width of a pencil, 6.00 mm?

(a) ________ Hz
(b) ________
(c) ________

1.5 The rather unappetizing confection you see in the photograph is a carrot strip that has been cooked in a microwave oven from which the rotating turntable was removed. When it is turned on, the oven uses a "magnetron" tube to generate standing microwaves in the cavity of the oven. No cooking heat is generated at the locations of the standing-wave nodes, and maximal heat (microwave "hot spots") is generated at the locations of the standing-wave antinodes. The manufacturer's label on the oven states that it operates at a frequency of 2450 MHz. (a) Based on this frequency, what is the wavelength of the microwaves generated inside the oven? (b) Based on your answer to part a, what is the distance between two antinodes of the standing microwaves? Give your answer to three significant figures. (c) Is the result in part b compatible with the measurement shown in the photograph?

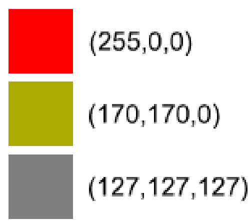


(a) __________ cm
(b) __________ cm
(c) ○ Yes ○ No

1.6 Apollo 11 astronauts left an array of "corner mirrors" on the Moon, each one consisting of three mirrors arranged at mutual right angles like the corner of a box. A corner mirror has the property that light shined on it from any angle is reflected back exactly in the direction from which it came. Scientists subsequently directed a brief pulse of light from a powerful laser at the mirror assembly on the Moon, and measured the time required for it to travel to the array and return back to Earth: 2.56 s. How far away was the array of mirrors at that time?

__________ m

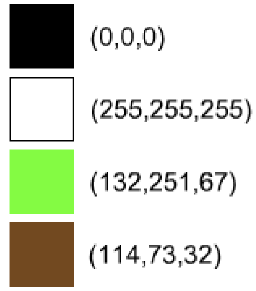
1.7 The human eye contains receptors for just three colors of light: red, green, and blue. All the colors we perceive in the world are mixed in our brains from different proportions of these three **primary colors**.
Computer and video display screens, both CRTs and flat panels, take advantage of this fact by using color-generating elements of the same three colors to create the appearance of any imaginable color. An image on the screen is built of tiny dots of color called picture elements, or "pixels". (By default, the period at the end of this sentence consists of just one black pixel.) Each pixel has three different color levels associated with it, one level for each of red, green, and blue, with integer values from 0 (representing no color), to 255 (representing the brightest level of a color). For example, (R, G, B) = (255, 0, 0) represents the brightest pure red, (170, 170, 0) represents an equal mixture of red and green, without any blue, which comes out a sort of dark yellow, and (127, 127, 127) represents an equal mixture of all three primary colors, each one at half its maximum intensity: a color sometimes called "50% gray".



(a) If black is the "color" corresponding to a lack of visible light, what are the RGB values of a black pixel? (b) If white is the "color" corresponding to the brightest possible equal mixture of all colors, what are the RGB values of a white pixel? (c) How many different colors can a pixel be? (The answer is large, but enter it as a whole number.)

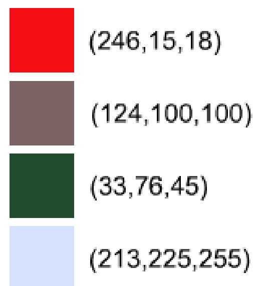
(a) (__________ , __________ , __________)
(b) (__________ , __________ , __________)
(c) __________ colors

1.8 The **brightness** of a computer screen color is a percentage value calculated by dividing the brightest of its three color levels by 255, the maximum possible brightness. For example, the brightness of the magenta hue (R, G, B) = (127, 0, 135) is 135/255 or 53%. What is the brightness of each of the following colors: (a) black (0, 0, 0); (b) brightest white (255, 255, 255); (c) chartreuse (132, 251, 67); (d) brown (114, 73, 32).



(a) ________ %
(b) ________ %
(c) ________ %
(d) ________ %

1.9 Equally balanced computer screen colors like (R, G, B) = (0, 0, 0), (255, 255, 255), or (120, 120, 120) have no predominant hue, and they look black, white, or some shade of gray in between: The grays with greater brightness are closer to white, and the grays with less brightness are closer to black. A color like (104, 100, 100) will appear slightly red, but it is mostly the gray color (100, 100, 100) with only a small additional amount of red. Color specialists say that this kind of color is "unsaturated".



Colors that are "saturated" have a rich predominant color with little gray mixed in. For example the RGB color (246, 15, 18) is a very saturated red. It only contains a slight constituent of gray, corresponding to its lowest-level component: (15, 15, 15). The **saturation** of an RGB color equals the percentage of its brightest component color that does **not** form part of its gray constituent. If the brightest component of an RGB color has value N, and the dimmest component has value n, then its gray constituent is (n, n, n) and its saturation equals $(N - n)/N$. For example, the saturation of the vibrant red mentioned above is (246 − 15)/246 = 94%. (a) What is the saturation of the red-tinged gray color (124, 100, 100)? (b) What is the saturation of the dark green (33, 76, 45)? (c) What is the saturation of the light blue (213, 225, 255)?

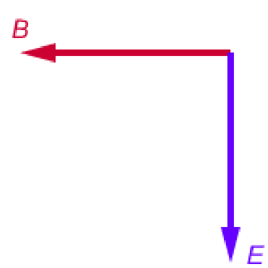
(a) ________ %
(b) ________ %
(c) ________ %

1.10 A live outdoor radio broadcast is disrupted when someone fires a gun into the air right in front of the microphone. The sound of the shot is broadcast via radio waves to a communications relay satellite in a geosynchronous orbit 35,790 km directly overhead, and back down to the radio in a police car parked 85.0 m away from the microphone. The sound waves of the gunshot also travel directly through the air to the car, at a speed of 343 m/s. Which sound does a police officer in the car hear first, the broadcast sound or the direct sound?

○ Broadcast ○ Direct

Section 2 - Electromagnetic waves

2.1 The diagram shows the electric and magnetic field components of an electromagnetic wave at a certain location and at a certain instant in time. Is the wave traveling toward you or away from you?



○ Toward you ○ Away from you

Section 3 - Proportionality of electric and magnetic fields

3.1 An electromagnetic wave in a vacuum has an electric field amplitude of 417 V/m. What is the magnetic field amplitude of this wave?

__________ T

3.2 An electromagnetic wave in a vacuum has an instantaneous magnetic field strength of 7.50e−7 T at a certain point. What is the instantaneous electric field strength of the wave at this point?

__________ V/m

3.3 Here is a pair of sinusoidal functions describing an electric and a magnetic field that vary in space and time.

$$E = \left(262\frac{\mathrm{V}}{\mathrm{m}}\right)\cos\left[\left(3.57\times 10^{7}\frac{\mathrm{rad}}{\mathrm{m}}\right)x - \left(1.07\times 10^{16}\frac{\mathrm{rad}}{\mathrm{s}}\right)t\right]$$

$$B = \left(8.74\times 10^{-6}\mathrm{T}\right)\cos\left[\left(3.57\times 10^{7}\frac{\mathrm{rad}}{\mathrm{m}}\right)x - \left(1.07\times 10^{16}\frac{\mathrm{rad}}{\mathrm{s}}\right)t\right]$$

They cannot be the field equations of an electromagnetic wave. Why not?

Section 6 - Electromagnetic energy: the Poynting vector

6.1 In the text it is stated that when an electromagnetic wave intersects a surface obliquely, the instantaneous area power density conveyed to the surface depends on the angle θ between its area vector and the Poynting vector of the wave. The equation expressing this relationship is

$$S_\theta = \mathbf{A}\cdot\mathbf{S}$$

where S_θ is the oblique area power density, $\mathbf{A}$ is the area vector, and $\mathbf{S}$ is the Poynting vector. Since $\mathbf{S}$ is itself a cross product, S_θ can be written as

$$S_\theta = \mathbf{A}\cdot\left(\frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}\right) = \frac{1}{\mu_0}\mathbf{A}\cdot(\mathbf{E}\times\mathbf{B})$$

The vector product $\mathbf{A}\cdot(\mathbf{E}\times\mathbf{B})$ is called a vector *triple product*.

In the chapter on rotational dynamics, a computational formula based on a *determinant* was given for the vector cross product:

$$\mathbf{E}\times\mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ E_x & E_y & E_z \\ B_x & B_y & B_z \end{vmatrix}$$

For this problem, show that a determinant formula for the vector triple product is

$$\mathbf{A}\cdot(\mathbf{E}\times\mathbf{B}) = \begin{vmatrix} A_x & A_y & A_z \\ E_x & E_y & E_z \\ B_x & B_y & B_z \end{vmatrix}$$

From this you can conclude that a general formula for the instantaneous area power density of an electromagnetic wave striking a surface at the oblique angle θ is

$$S_\theta = \frac{1}{\mu_0}\begin{vmatrix} A_x & A_y & A_z \\ E_x & E_y & E_z \\ B_x & B_y & B_z \end{vmatrix}$$

Section 7 - Intensity and energy density

7.1 An electromagnetic wave has an electric field amplitude of 320 V/m. (a) What is the maximum magnitude of the Poynting vector of this wave? (b) What is the average magnitude of the Poynting vector? (c) What is the intensity of the wave? (d) What is the root mean square of the electric field of the wave?

(a) __________ $\mathrm{W/m^2}$
(b) __________ $\mathrm{W/m^2}$
(c) __________ $\mathrm{W/m^2}$
(d) __________ V/m

7.2 An electromagnetic wave in a vacuum has a maximum magnetic field amplitude of 5.67×10^{-7} T. (a) What is the maximum electric field amplitude of the wave? (b) What is the average power per unit area associated with the wave?

(a) __________ V/m
(b) __________ W/m^2

7.3 An electromagnetic wave has an electric field strength of 6.00×10^4 V/m and a magnetic field amplitude of 2.00×10^{-4} T at a certain location. (a) What is the electric field energy density at this location? (b) What is the magnetic field energy density? (c) What is the total energy density?

(a) __________ J/m^3
(b) __________ J/m^3
(c) __________ J/m^3

7.4 What is the average total energy density in electromagnetic radiation that has an intensity of 475 W/m^2?

__________ J/m^3

7.5 A neodymium laser emits a pulse of radiation, providing 125 TW of power for a duration of 1.00 ns. How much energy is contained in that pulse?

__________ J

7.6 An electromagnetic wave has an electric field strength of 145 V/m at a point P in space at time *t*. (a) What is the electric field energy density at P? (b) A parallel-plate capacitor whose plates have area 0.106 m^2 has a uniform electric field between its plates with the same energy density as the answer to part a. What is the charge on the capacitor?

(a) __________ J/m^3
(b) __________ C

7.7 An electromagnetic wave has an magnetic field strength of 6.78×10^{-6} T at a point P in space at time *t*. (a) What is the magnetic field energy density at P? (b) A 0.410 m long solenoid consisting of 950 loops of wire has a uniform magnetic field inside its coil with the same energy density as the answer to part a. The radius of the coil is 0.215 cm. What is the current passing through the solenoid?

(a) __________ J/m^3
(b) __________ A

7.8 The instantaneous electric field energy density of an electromagnetic wave is $u_E = \varepsilon_0 E^2/2$. The instantaneous magnetic field energy density of an electromagnetic wave is $u_B = B^2/2\mu_0$. Show that $u_E = u_B$ at any instant in time.

7.9 The instantaneous electric field energy density of an electromagnetic wave is $u_E = \varepsilon_0 E^2/2$. The instantaneous magnetic field energy density of an electromagnetic wave is $u_B = B^2/2\mu_0$. (a) What is the average electric field energy density over time, expressed in terms of the electric field amplitude, E_{max}? (b) What is the average magnetic field energy density over time, expressed in terms of the magnetic field amplitude B_{max}? (c) Show that the average total energy density of the wave is $u_{avg} = \varepsilon_0 E^2_{max}/2$.

Section 10 - Radiation intensity and distance

10.1 The SMART-1 lunar orbiter, launched by the European Space Agency, reached the Moon in November 2004 after a 13-month voyage from the Earth. It used an ion propulsion system powered by two winglike solar panel assemblies, one of which you see deployed for testing in the photograph. The assembly consists of lithium photovoltaic cells and it measures 14 m wide by 115 cm tall. In flight, when the panels are adjusted to directly face the Sun, each assembly produces 1.9 kW of power. It was shown in the text that the intensity of sunlight is 1380 W/m^2 at the distance of the Earth's (and the Moon's) orbit. What is the efficiency, expressed as a percentage, of the SMART-1 solar panels at converting sunlight power into electric power?

__________ %

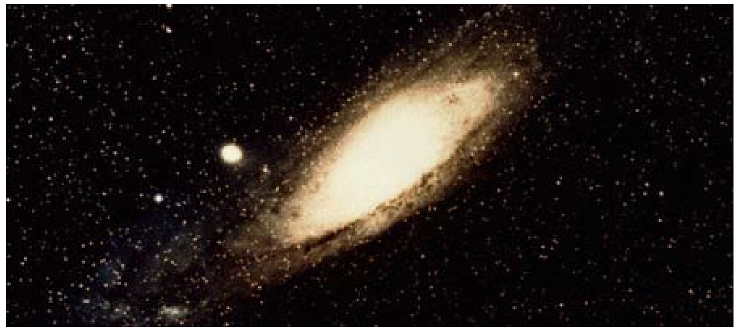
10.2 Sunlight falling on the upper layers of the Earth's atmosphere has an intensity of 1380 W/m^2. If all the energy were carried by a single wave, (a) what would be the electric field amplitude E_{max} for this wave? (b) What would be the magnetic field amplitude B_{max}?

(a) __________ V/m
(b) __________ T

10.3 Suppose a blue-giant star is located 10.7 light-years from the solar system and radiates energy at a rate 135 times that of the Sun. What would be the intensity of the starlight reaching the Earth from this star? (Treat this as an ideal problem, and ignore any absorption due to interstellar dust or the Earth's atmosphere.)

__________ W/m^2

10.4 The Andromeda galaxy pictured here is a vast whirl of 1.0×10^{11} stars, located 2.9 million light-years from the Earth. Assume that, on average, each star in this galaxy emits radiation with the same power as the Sun. (a) What is the intensity of light from this galaxy as it reaches the Earth? (Treat this as an ideal problem and ignore any absorption due to interstellar dust or to the Earth's atmosphere.) (b) A candle emits radiation in all directions with a power of 4.2 W. What is the intensity of the candlelight at a distance of 9.3 km? (c) From which source is the light more intense, the galaxy or the candle?



(a) __________ W/m^2
(b) __________ W/m^2
(c) ○ The galaxy ○ The candle

10.5 The Voyager spacecraft was launched in 1977 on a mission to explore Jupiter, Saturn, and the reaches of interstellar space beyond them. It is by far the most distant human-made object in the universe: In September 2004 this craft was 14 billion kilometers from the Earth. Voyager uses a directional antenna that broadcasts radiation in a solid angular swath that covers 1.2% of its sky, with a power of 23 W. What was the intensity of its signal at the Earth in September 2004? (Find the maximum possible intensity by ignoring the attenuating effects of the Earth's atmosphere and interstellar dust.)

__________ W/m^2

10.6 The radio telescope in Arecibo, Puerto Rico has a diameter of 300 m, and is able to detect radio signals with a power of as little as 6×10^{-22} W. If the alien inhabitants of a planet 100 light-years away are trying to contact us, what is the minimum power their transmitter must have if its signal is to be detectable at Arecibo? Assume that the aliens transmit omnidirectionally. State your answer to two significant digits.

__________ W

10.7 A tiny source is emitting electromagnetic radiation equally in all directions. At a distance of 1.0 cm this radiation can be considered to be a plane wave over a small enough region. At a point P 1.0 cm from the source, the electric field amplitude of the wave is 3.5 V/m. (a) What is the amplitude of the magnetic field of the wave at P? (b) What is the average intensity of the radiation at P? (c) What is the power with which the source emits radiation?

(a) __________ T
(b) __________ W/m^2
(c) __________ W

10.8 A 40 W electric light bulb emits of 3.25% of its consumed energy as blue light with a wavelength of 433 nm. Assume this energy is contained in a single electromagnetic wave, and write the plane wave field equations that approximately describe the light from this bulb at a distance of 12.5 m.

Section 12 - Intensity and field strength around a dipole antenna

12.1 Due to its high electrical conductivity, seawater weakens ordinary radio waves rapidly as they pass through it. Extremely low frequency (ELF) radio waves in the range from 40−80 Hz are the least subject of all radio waves to this kind of attenuation, and so are an attractive option for communicating with cruising submarines. Signals generated at a land-based naval station are broadcast toward the sky, where they bounce off the ionosphere (an electrically conductive layer high in the atmosphere), and then propagate downward to penetrate the sea and reach their recipients.
One of the difficulties associated with implementing ELF is the problem of generating a signal in the required frequency range. (a) What is the wavelength of an ELF radio wave having a frequency of 75 Hz? (b) How long is a half-wave dipole antenna that broadcasts waves of this frequency? (c) An AC-driven linear antenna with the power feed at one end rather than in the middle is called a quarter-wave dipole. How long is a quarter-wave dipole antenna that broadcasts radio waves with a frequency of 75 Hz?
(The longest ELF antenna ever constructed is a horizontal pole-mounted wire constructed by the U.S. Navy in Wisconsin, 222 km long. It is neither a half-wave nor a quarter-wave dipole, and uses technological refinements to achieve signals in the desired frequency range.)

(a) __________ m
(b) __________ m
(c) __________ m

Section 13 - Radiation pressure

13.1 A regulation NFL football field, including its end zones, is exactly 120 yards long and 160 feet wide. At noon on a certain cloudless day, the Sun is shining on the field with an intensity of 825 W/m^2. The albedo of the field is 0.550. (a) What is the downward force of sunlight on the field? (b) A newly minted half-dollar coin weighs 11.340 g. Which force is greater: that exerted by the light pressure on a football field, or the weight of two half dollars still sitting on it after several coin tosses by a clumsy referee?

(a) __________ N
(b) i. The force due to light pressure
ii. The weight of the coins

13.2 A **laser pointer** directs a narrow beam of light that can be used by a speaker to illuminate locations on a projected image during a lecture. If a 3.7 mW laser pointer illuminates a circular area with a radius of 1.2 mm on a projection screen with an albedo of 0.86, find the pressure the laser exerts on the illuminated portion of the screen.

__________ N/m^2

13.3 (a) What is the radiation pressure that a 250 W light bulb emitting light with an efficiency of 15% exerts on a small matte-black surface at a distance of 0.500 m? (b) What is the pressure it exerts on a small mirror at the same distance?

(a) __________ N/m^2
(b) __________ N/m^2

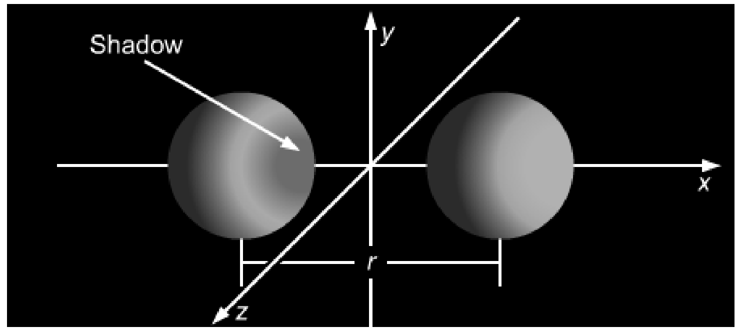
13.4 The Echo 2 satellite, a perfectly reflective aluminized Mylar® sphere 41.0 meters in diameter, was launched in 1964 into an orbit a little more than 1000 km above the Earth's surface. It was an experimental "passive communications" satellite, serving to bounce radio signals from a ground transmitter to a receiver a quarter of the globe away. As the first human-made object visible from space, it inspired countless young future physicists with its visible progress as a bright star drifting across the heavens. (a) The distance of Echo 2 from the Sun was 1.50×10^{11} m, and assume the the power emitted by the Sun is 3.91×10^{26} W. What was the intensity of the sunlight striking the satellite? (b) What pressure did the sunlight exert on Echo 2? (c) What force did sunlight exert on the satellite? (Hint: For the purposes of this calculation, you may consider the sphere to be a flat disk with a diameter of 41.0 m, directly facing the Sun.) (d) The mass of the Echo 2 balloon was 256 kg. What was the magnitude of its acceleration due to the pressure of sunlight?

(a) __________ W/m^2
(b) __________ N/m^2
(c) __________ N
(d) __________ m/s^2

13.5 (a) The Sun repels the Earth by pushing on it with sunlight. What is the force exerted by sunlight pressure on the planet Earth? Treat the Earth as a flat disk facing the Sun. Its radius is 6.37×10^{6} m, and its albedo is 0.39. The intensity of sunlight at the Earth's orbit is 1380 W/m^2. (b) The Sun attracts the Earth by pulling on it with gravity. What is the force exerted by the Sun's gravity on the planet Earth? The mass of the Earth is 5.97×10^{24} kg, the mass of the Sun is 1.99×10^{30} kg, and the distance between the two bodies is 1.50×10^{11} m.

(a) __________ N
(b) __________ N

13.6 Clouds of interstellar dust and gas coalesce into nodes, which further contract into protostars that eventually give birth to new stars. The mutual gravitational attraction of dust particles in the cloud accounts for some of this attraction, and it becomes the dominant factor as the cloud gathers into an opaque mass. During the early stages of contraction, however, the inward light pressure of the entire surrounding universe on the dust cloud is the dominant factor. How does this compressing effect vary with the size of the dust cloud, which is to say with the distances between its particles? In this problem you will answer this question by considering how isotropic light pressure provides an "attractive" force between two perfectly reflective dust particles.



(a) Imagine a three-dimensional *xyz* coordinate system in interstellar space. Two spherical dust particles of fixed diameter lie on the *x* axis at a variable distance *r* from each other, one on either side of the origin. Consider the sum of the light from all the stars with positive *x* coordinates that illuminates these particles. Which particle does this light repel with greater force? Explain your answer.
(b) The "positive *x*" particle casts a diffuse shadow on the "negative *x*" particle. The effective area of the shadow is proportional to the fraction of the hemispherical "positive *x*" sky that the "positive *x*" particle obscures, from the viewpoint of the "negative *x*" particle. What power of the distance *r* between the particles is this obscured fraction proportional to?
(c) The "positive *x*" sky repels both particles, but it repels the "negative *x*" particle less than it does the "positive *x*" particle because of the shadow. The difference between these repulsive forces is equivalent to an attractive force between the particles in their own frame of reference. Let *I* be the intensity of the net *x* component of all the starlight from the "positive *x*" sky, and let *A* be the effective area of the shadow that the "positive *x*" particle casts on the "negative *x*" particle. What is the magnitude *F* of the resulting "attractive force" between the particles?
(d) Symmetrically, the "negative *x*" particle shadows the "positive *x*" particle from the "negative *x*" sky, doubling the light pressure "attraction" between the particles. Considering your answers to parts b and c, what kind of proportionality relates the total "attractive force" between the particles and the distance *r* between them?

(a) ○ The "positive *x*" particle ○ The "negative *x*" particle
(b) ○ r^2 ○ r ○ $1/r$ ○ $1/r^2$
(c) ○ $F = 2I/c$ ○ $F = I/c$ ○ $F = 2IA/c$ ○ $F = IA/c$
(d) i. Direct square
ii. Direct
iii. Inverse
iv. Inverse square

13.7 A comet is a huge "dirty snowball" that orbits the Sun in an extremely long, eccentric orbit that loops close around the Sun at one end and travels far out into space on the other. Each time it nears a close encounter with the Sun, it is heated so that copious amounts of dust and gas boil off its surface (resulting in a lifetime that is quite limited in astronomical terms). Dust particles that boil off the comet do not follow on in its orbit. Instead, they are blown away from the Sun by the pressure of sunlight. Assume that each dust particle is spherical, with radius *r*, albedo 0.37, and density 3.90×10^3 kg/m^3. What is the radius *r* of a dust particle that is equally repelled by the Sun's light and attracted by its gravity?

__________ m

Section 18 - Polarization and intensity

18.1 A beam of unpolarized light of intensity 30.0 W/m^2 passes through a polarizing sheet. What is the electric field amplitude of the transmitted beam?

__________ V/m

18.2 This problem deals with combinations of ideal polarizing filters that have the net effect, first of polarizing incident light, and then of rotating the plane of polarization by 90°. In the text it was shown that two crossed polarizing filters (at 90° to each other) completely block initially unpolarized light that falls on the pair of filters. The first filter linearly polarizes the incoming light, resulting in an initially polarized intensity of I_p. Then the second crossed filter blocks all of the now-polarized light, and 0% intensity is transmitted. Somewhat surprisingly, adding a third filter between the crossed filters, so that each polarizer is rotated 45° with respect to the one just before it, does transmit some light. Again, the first filter linearly polarizes the incoming light, but the other filters cumulatively rotate this plane of polarization by 90° without blocking all of the light. In this case $I_{final} = 0.25I_p$, so 25.0% of the intensity emerging from the initial polarizer is transmitted through all three filters.

A combination of four filters, each rotated 30.0° clockwise with respect to the previous one, is exposed to light. This combination results in a 90° clockwise rotation of the plane of polarization that emerges from the first of the filters. What fraction of this initially polarized intensity I_p is the final intensity of the polarized light that emerges after passing through the other filters? Express your answer in terms of a **percentage** of I_p.

__________ %

18.3 Before working this problem, read as introductory background the previous homework problem which dealt with combinations of two, three, and four polarizing filters that have the net effect of polarizing incident light, and then rotating the plane of polarization by 90°.

A combination of 91 ideal polarizing filters, each rotated 1.00° clockwise with respect to the previous one, is exposed to unpolarized light. This combination results in a 90° clockwise rotation of the initial plane of polarization. The first filter linearly polarizes the incoming light. Call this initially polarized intensity I_p. What fraction of I_p is the final intensity of the polarized light that emerges after passing through all the filters? Express your answer in terms of the **percentage** of the intensity transmitted.

__________ %

Section 22 - Optically active substances

22.1 The organic chemical **lactic acid** occurs in two mirror image forms. D-lactic acid (which rotates polarized light clockwise) is produced by muscles as they work, and its excessive buildup is the fatigue poison that causes discomfort in overworked limbs. L-lactic acid (which rotates polarized light counterclockwise) is an ingredient of milk. A scientist wishes to establish an extremely accurate value for the specific rotation α_0 of lactic acid. She obtains a sample of pure L-lactic acid from milk, and then dilutes it with optically inactive water to a concentration of 325.0 kg/m^3. She measures the amount of counterclockwise rotation in the plane of 589 nm polarized light from a sodium lamp that results after the light passes through a 10.00 m long tube filled with the solution, and finds that it is 81.37°. (She knows the approximate specific rotation of lactic acid, so she is sure that the rotation really is this amount, and not the same angle plus some multiple of 180°.) What is the specific rotation of lactic acid, to four significant figures?

__________ ° m^2 /kg

22.2 Molecules of the organic terpene **limonene** occur naturally in mirror image forms, called D-limonene and L-limonene depending on whether they rotate polarized light clockwise (D) or counterclockwise (L). D-limonene is the essential oil that conveys the scent of orange, while L-limonene carries the scent of lemon. L-limonene is frequently used to add lemon scent to household products such as dish soaps and oven cleaners. The specific rotation of limonene is α_0 = 120 °m^2/kg.

L-limonene occurs naturally in lemon peels, but it is also a major constituent of a far cheaper byproduct of the timber industry, pine oil, where it occurs mixed with substantial amounts of D-limonene. Limonene refined from crushed pine needles is a variable mixture of both forms of the molecule, and there is no economical way to separate them. Perfume manufacturers wish to make sure that the expensive limonene they purchase for use in scents is pure L-limonene, uncontaminated by D-limonene. A scientist uses optically inactive alcohol to dilute a solution of limonene to a concentration of 125 kg/m^3. Testing it with a polarimeter, he finds that a sample of the solution 15.6 cm thick rotates polarized light 21.5° counterclockwise. What percent of the limonene is the desired form, L-limonene?

__________ %

chapter 35 Reflection

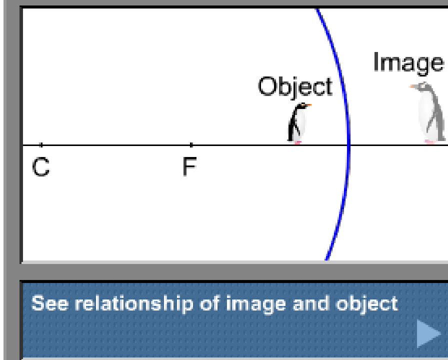
35.0 - Introduction

"Image is everything," one well-known advertisement blared for years. (An ad from another company then trumpeted the contradictory message that "image is nothing." Go figure!)

For physicists, image is as much a matter of mirrors and lenses as it is of appearance and athleticism. By arranging mirrors and lenses, they can magnify or shrink images and place them where they are needed. The design of a wide range of devices, from reading glasses and cameras to telescopes, depends on a thorough understanding of how to manipulate images created with light.

This chapter begins with a discussion of light rays and mirrors, two essential elements of your understanding of images. The simulation to the right features a concave mirror. The mirror creates images of objects. In this case, the object is a penguin. You can move the penguin left and right, and observe how its positioning changes the location and size of the image created by the mirror. The image is shown as a faded out version of the object itself. At some locations of the object, the image will be off the screen, but it will reappear when you drag the object to another location.

interactive 1



As you move the object back and forth, consider the following questions. Where can you place the object so that the image and object are on the same side of the mirror? On opposite sides? Where can you place the object so that the image is smaller than the object? Larger? The variety of images produced by different types of mirrors is a major topic in this chapter.

If you press the SHOW RAYS button in the simulation, you will see three light rays that emanate from the penguin, reflect off the mirror, and converge to define the top of the image. These rays are used extensively in the study of mirrors and lenses, and this is your chance to begin to experience their properties. (The rays do not always converge perfectly; this is a realistic depiction of the way curved mirrors work.) You can turn off the rays by pressing the HIDE RAYS button.

35.1 - Light and reflection

Reflection: Light "bouncing back" from a surface.

To form a mirror image, light bounces back from the mirror's surface.

When you look at yourself in a mirror, you are seeing a reflection of yourself. When you look at the Moon at night, you are seeing sunlight reflecting off that distant body.

Not all the light that reaches a surface reflects. In fact, you see an object like a tree as having different colors because its varied parts reflect some wavelengths of light and absorb others. Light can pass through a material, as it does with a glass window. It can also be absorbed by a material, as evidenced by how a black rock warms up during a sunny day. All this can happen simultaneously: Light will reflect off the surface of a lake (which is why you see the lake), penetrate the water (otherwise, it would be completely dark below the surface), and be absorbed by the water, warming it.

concept 1

Reflection
Light bouncing back from a surface

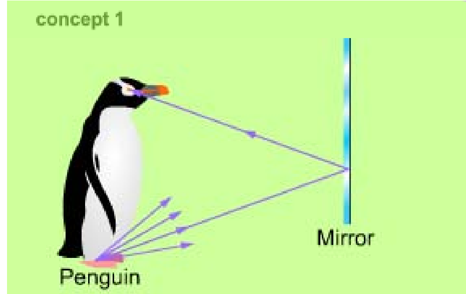
To understand reflection, it is often useful to treat light as a stream of particles that move in a straight line and change direction only when they encounter a surface. Each light "particle" acts like a ball bouncing off of a surface, and like a ball, it reflects off the surface at a rebound angle equal to its incoming angle. You see yourself in a mirror because the light bounces back to your eyes from the mirror.

The term "reflection" likely conjures up images of light and perhaps mirrors. Studying mirrors is a good way to learn about reflection because they are designed to reflect light in a way that creates a clear visual image. However, it is worth noting that reflection does not apply only to light. Some creatures use the reflection of sound (echoes) to help them perceive their surroundings and stalk their prey. For example, bats, seals and dolphins emit high frequency sound and then listen for the reflected waves. By analyzing these reflections, they can "see" with great precision.

Radar, used to track airplanes, is based on the reflection of radio waves. A sophisticated understanding of reflection can be used to design "stealth" aircraft that are difficult to detect with radar. Stealth aircraft register on radar screens as being about as large as a BB, in part because of their ability to reflect incoming waves in "random" directions.

35.2 - Light rays

Light ray: A straight line used to represent the path of light.



Light rays
Represent light with straight lines
Used to analyze mirrors, lenses

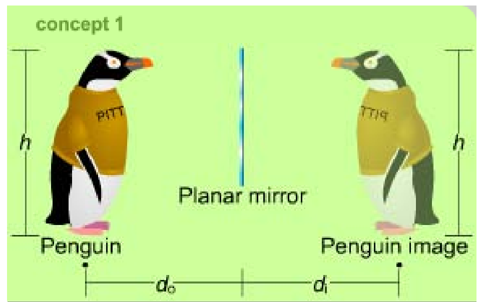
Light rays are used to analyze devices such as mirrors and lenses. Light emanating from a point in a particular direction is treated as moving in a straight line. You can think of this as modeling light with a collection of laser beams, each ray moving in a straight and narrow line until it intersects a surface.

You can see an example of light rays depicted on the right, where a penguin views himself in a mirror. Several rays are shown emanating from a point on the penguin's foot. We show the full path of a single ray to illustrate how the penguin can see his foot reflected in the mirror. The light ray starts at the foot, reflects at the mirror and then reaches his eye. The arrowheads indicate the direction of the light's travel.

The penguin's feet are not the source of the light. That might be the Sun or an electric lamp. But for the purposes of analyzing the mirror, we will proceed as though the light rays originate from the object whose image we intend to analyze.

Light is a complex phenomenon. Treating it as propagating in a straight line does not explain why a beam from a flashlight spreads out as it travels; other concepts are needed there. But to understand mirrors and lenses, light rays prove very useful.

35.3 - Mirror basics



Planar mirror object and image
Same height
Equally far from mirror
Image is right side up

At the right, you see a penguin looking at himself in a *planar mirror*. A planar mirror is a flat, smooth surface that reflects light. A typical planar mirror is manufactured by coating one side of a sheet of glass with a metallic film. It is designed to reflect as much light as possible and to create distinct images. Either by looking at yourself in a planar mirror or referring to the illustration in Concept 1, you can observe four essential properties of an image produced by this kind of mirror:

1. Your image is the same size as you are. The variable h is used to represent its height.
2. Your image appears as far behind the mirror as you are in front of it. Your distance is called the object distance, and how far the image is behind the mirror is called the image distance. The object distance, represented by d_o, is positive by convention. When an image is on the opposite side of a mirror from the object, its distance d_i from the mirror is negative by convention.
3. The image has front-back reversal. If you are facing north, then your mirror image is facing south. If this reversal did not occur, you would see your back.
4. The image is **not** reversed either right-left or up-down. If you look in a mirror and raise your left hand, its reflection will rise up on your left. Your head also still appears on the top of the image. The image is not inverted.

Mirrors do create the illusion of left-right reversal. For instance, if you raise your left hand, it will appear on your left but on the image's "right," because your image is facing the opposite direction than you are. This is akin to "stage left" or "stage right" directions for actors, which are based on the perception of what is left or right for the audience, not for the actors themselves as they face the audience. Their "left" and "right" are reversed because they are facing in the opposite direction than the audience, just as your mirror image is.

You cannot easily read normal writing in a mirror (unless you are as gifted as Leonardo da Vinci, who wrote in "mirror writing" so that others could not easily decipher his work). Mirror writing does have its uses: An emergency vehicle may have "ɘɔnɒludmA" written on its front so that the text can be readily read when seen in the rearview mirror of a car.

Image: A reproduction of an object by means of light.

Virtual image: An image on the opposite side of a mirror than the object it is created from. The image cannot be projected onto paper at its perceived location.

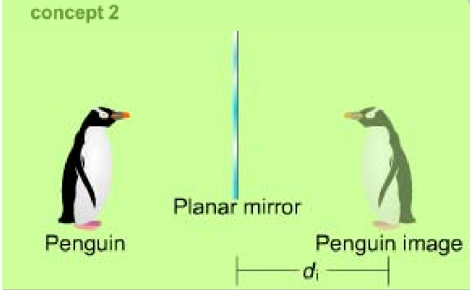
Kitten looking for its image behind a mirror.

Real image: An image on the same side of a mirror as the object it is created from. The image can be projected onto paper at its perceived location.

concept 1

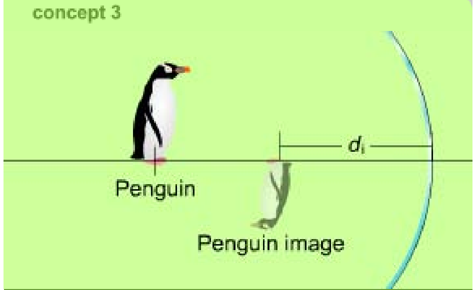
Image
A reproduction created by light

concept 2



Virtual image
Opposite side of mirror
Negative distance

concept 3



Real image
Same side of mirror
Positive distance

When you look into a planar mirror, your image appears to be on the **opposite** side of the mirror, in a "looking glass" world that you cannot reach except in literature. The light of the image registered by your eyes is not actually coming from behind the mirror, even though your brain may interpret it this way.

You may be so accustomed to virtual images such as those formed by planar mirrors that you take them for granted and believe them to be as "real" as any other image. But not as real as any other **object**: A true understanding of virtual images must be learned. A kitten, when first exposed to a mirror, believes its reflected image to be another kitten and will search for the potential playmate behind the mirror. Over time, some animals (including humans, chimpanzees and dolphins) can learn how to interpret the nature of the images created by mirrors.

It will now be helpful to be more formal about some basic terminology. Objects are "in front" of a mirror: They are on its reflecting side. (In the diagrams in this book we will almost always place an object to the left of the mirror, but that certainly is not a law of physics, just a convention of sorts about which side is "in front.") If the image is on the same side of the mirror as the object, it also is in front of the mirror. If the image is on the opposite side of the mirror from the object, we say it is "behind" or in back of the mirror. A virtual image created by a mirror is behind the mirror. The distance between a virtual image and the mirror is negative.

Curved mirrors can create what is called a real image. In Concept 3, you see a diagram of a concave mirror creating a real image: an image on the same side of the mirror as the object. By convention, the distance between the mirror and the real image is positive. We stress the conventions of sign because they will prove essential when equations are used to calculate certain properties of an image, such as its position.

The terms "real" and "virtual" can be confusing. One way to decide whether an image is real or virtual is to imagine placing a piece of paper at the location of the image and then observing the results. With real images, the light passes through the location of the image; with virtual images, it does not. If you place a piece of paper "behind" a planar mirror, where the virtual image is located, you will not find an image projected there. With virtual images the brain projects the location of the image as it interprets the light it receives.

On the other hand, if you place a piece of paper at the location of a real image, you will see the image on the paper. For instance, images created by movie projectors are real, which is why you can see them on movie screens.

Concept 4 shows an optical illusion created by a popular toy. A special combination of curved mirrors creates a real image that is so solid looking you are tempted to reach out and grasp it.

concept 4

An optical illusion
Real image perceived as object

35.5 - The law of reflection

Law of reflection: The angle of incidence equals the angle of reflection.

Diffuse and specular reflections from holiday ornaments.

At the right, you see a planar mirror. A light ray from the penguin's foot is reflecting off the mirror. The light ray before reflection is called the *incident ray*; the light ray that reflects off the mirror is known as the *reflected ray*. In this illustration, the rays are in the same vertical plane.

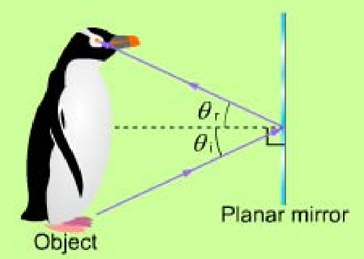
The *angle of incidence* θ_i is shown in the diagrams to the right. The angle θ_i represents the angle between a line perpendicular (normal) to the mirror's surface and the incident ray of light.

The light leaves the surface at the *angle of reflection* θ_r. This is the angle between the normal line and the reflected ray. As you can see in the diagrams, the angles of incidence and reflection are the same. This is the law of reflection: The angle of incidence equals the angle of reflection. This law is confirmed by experiments and theory.

The law of reflection applies to smooth surfaces, which exhibit *specular reflection*. The right-hand ornament in the photograph above is extremely smooth and produces a specular reflection. Rays that are parallel and close together when they strike the ornament will all be moving in a new direction after they reflect, but they will still be parallel. Specular reflection is the topic of this chapter.

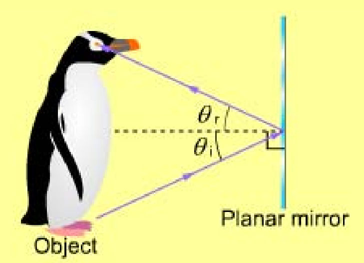
If the reflecting surface is rough, *diffuse reflection* results. The left-hand ornament above provides an example of diffuse reflection. With a rough surface, neighboring incident rays will reflect in a variety of directions. Rays that are parallel and close together when they strike the surface will not be parallel after they reflect. Diffuse reflection is sometimes desirable. For instance, matte (low gloss) wall paint is designed to achieve diffuse reflection for surfaces where a shiny appearance is undesirable.

concept 1



The law of reflection
Incidence angle equals reflection angle
Angles measured between light rays and normal line

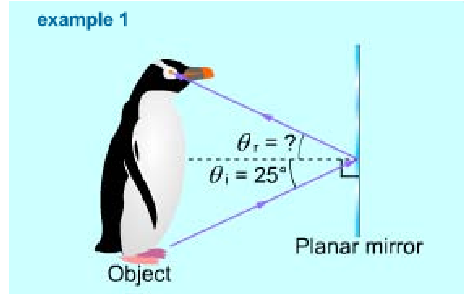
equation 1



The law of reflection

$$\theta_i = \theta_r$$

θ_i = angle of incidence
θ_r = angle of reflection



What is the angle of reflection?
Same as angle of incidence: 25°

35.6 - Ray diagrams for planar mirrors

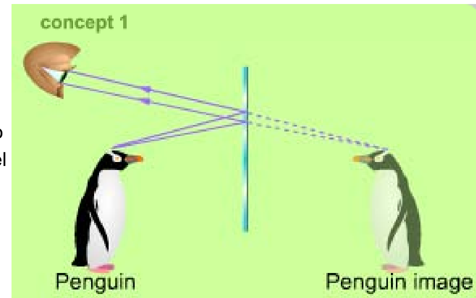
In this section, we use ray diagrams to explore some properties of an image produced by a planar mirror. To review these: The image created by a planar mirror is virtual, upright (right side up), the same distance from the mirror as the object, and the same size as the object.

Using the ray diagram in Concept 1, we first show why the image is virtual. You see two rays emanating from a point on the penguin's head. They reflect off the mirror and travel to the eye of an observer viewing the penguin's image. In this diagram, we are not concerned with what the penguin sees. We are using him as an object whose image is being viewed by an observer.

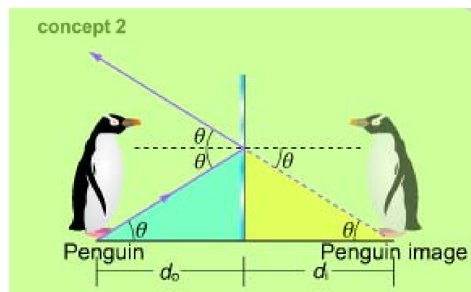
Because the observer's brain assumes that light travels in a straight line, it projects an image behind the mirror. Dashed lines, called *virtual rays*, show the paths the brain presumes are followed by rays emanating from the source of the light. The brain locates the top of the image at the point where these virtual rays converge. Since this point is behind the mirror, and no light actually comes from it, it forms part of a virtual image.

Rays also emanate from a point on the penguin's foot and reflect from a lower point on the mirror back to the observer's eye. The observer's brain locates the bottom of the image at a point where the virtual extensions of those rays intersect, a point behind the mirror and below the location of the image's head. This confirms that the image is not only virtual, but upright as well.

The diagram in Concept 2 shows that the image is the same distance behind the mirror as the object is in front. The two colored triangles share a common side (the mirror) and two pairs of equal angles, including a pair of right angles. This means the triangles are congruent, so their bottom legs must have equal lengths, and d_i and d_o must have equal magnitudes. A similar argument with two other congruent triangles demonstrates that the image has the same height as the object.



Ray diagram
Draw rays using law of reflection
Find image location with virtual rays



To relate object, image distances
Draw a ray
Identify equal angles
Congruent triangles prove object, image distances have same magnitude

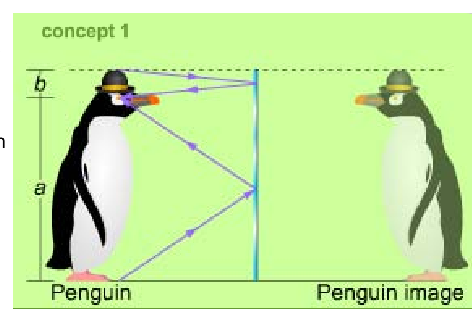
35.7 - How tall should a full-length mirror be?

To provide a chance for you to practice your understanding of planar mirrors, we will step through two ray diagrams that are used to answer a single question: How tall must a "full-length" mirror be, relative to a penguin, so that he can see himself from hat to foot?

The diagram for Concept 1 shows a penguin looking at himself in a mirror. We start with a mirror that is the same height as the penguin plus his hat, with the top of the mirror even with the crown of the hat, and the bottom of the mirror even with the soles of his feet.

Two reflecting light rays are shown. One starts at the penguin's foot, reflects off the mirror, and travels to his eye. The other starts at the top of his hat, and reflects back to reach his eye.

We will show that the mirror must be half the height of the penguin (and hat) in order for him to view himself in all his sartorial splendor. First, we will consider the foot-to-eye ray in more detail, using the illustration in Concept 2.

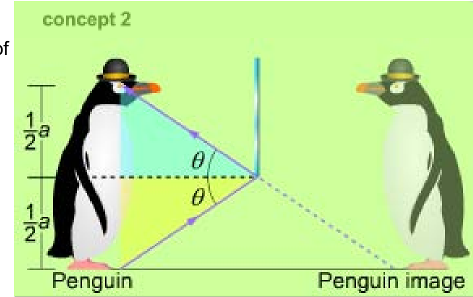


The reflecting penguin
How much of the mirror is needed for the penguin to see all of himself?

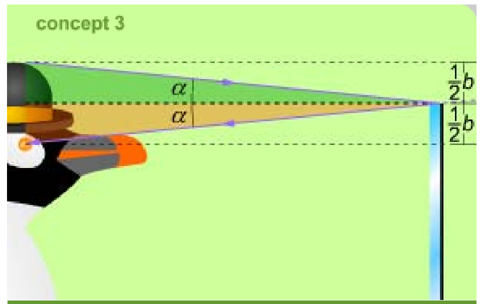
The law of reflection states that the angles of incidence and reflection must be the same, so the two angles labeled θ in the illustration are equal. The yellow triangle is formed with the incident ray as one side, the normal to the mirror as another, and part of the penguin as the third side. The sides of the blue triangle are the reflected ray, the normal, and part of the penguin. The distance from the ground to the penguin's eye we call a.

1. The blue and yellow triangles share a common side, the normal to the mirror. They also have two pairs of equal angles: two right angles where the normal line meets the penguin, and the two angles θ. This means the triangles are congruent.
2. The penguin can see his knees and waist in the mirror. Rays reflecting from these portions of his body to his eye will strike the mirror at points **higher** than the foot-to-eye reflection point. This means he does not need the part of the mirror below this point. Because the two triangles are congruent, the discarded portion of the mirror has height ½ a.
3. We call the distance from the penguin's eye to the crown of his hat b. Using the same reasoning as above, as illustrated in Concept 3, we find that the portion of the mirror of height ½ b at the top is not needed for him to see the crown-to-eye portion of his image.

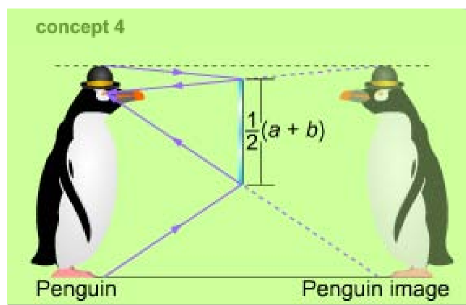
We started with a mirror the height of the penguin plus his hat, $a + b$. We discarded ½ a at the bottom of the mirror and ½ b at the top. We have shown that the required mirror height, $\frac{1}{2}(a + b)$, is half the penguin's total height. This result does not depend on the object and image distances, that is, on how far the penguin is standing from the mirror.



Find lowest point required
Draw reflected ray from foot to eye
$\theta_i = \theta_r$ so triangles congruent
Bottom part of mirror not needed



Find highest point required
Draw reflected ray from crown to eye
$\alpha_i = \alpha_r$ so triangles congruent
Top part of mirror not needed



Finding the minimum mirror height
Mirror is half-height

35.8 - Spherical mirrors

Spherical mirror: A mirror that is a portion of a sphere.

Concave mirror: A curved mirror whose reflective side is on the inside of the curve.

Convex mirror: A curved mirror whose reflective side is on the outside of the curve.

Concave and convex spherical mirrors.

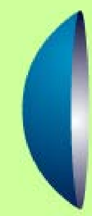
Mirrors can be curved as well as flat. One type of curved mirror is a spherical mirror, which consists of a section of a sphere. To describe spherical mirrors it is common to consider a sphere with silvered reflecting surfaces. In a concave mirror, reflection occurs on the inside surface, as illustrated in Concept 2. A convex mirror is made by using a section of the exterior of the sphere, as shown in Concept 3.

If you like mnemonics, a good way to remember which type is which is to recall that a con**cave** mirror is like a **cave** whose walls curve around you. Though we concentrate on spherical mirrors here, concave and convex are terms that can apply to curved mirrors other than spherical mirrors.

The photos above show two examples of spherical mirrors and the images they create. The left-hand photo shows a concave makeup mirror. These mirrors, often found in bathrooms, are designed to magnify an image: You look in the mirror and see a larger image of yourself. Lucky you!

The photograph on the right above shows a convex safety mirror that allows drivers to see around a corner of a twisty seaside road. Security mirrors are also often convex. They provide a larger field of view than a flat mirror, enabling store personnel, for example, to survey a large area. Convex auxiliary mirrors are often affixed to the flat rearview mirrors of trucks and RVs.

concept 1

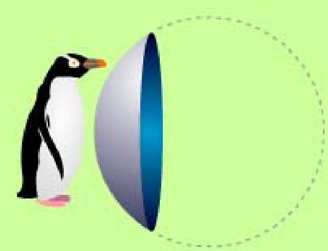


Spherical mirror
A portion of a sphere

concept 2

Concave mirrors
Reflecting surface is interior of sphere

concept 3



Convex mirrors
Reflecting surface is exterior of sphere

35.9 - Mirror terminology

Before we proceed to a further discussion of mirrors, we need to introduce some terminology. The diagrams to the right are used to illustrate these new terms. Many of the same terms are used to describe lenses, a topic you will study later.

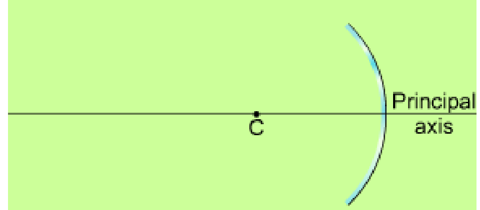
The mirrors discussed in this section are spherical. The sphere's *center of curvature* is shown as point C in the diagrams.

A line passing through the center of curvature and the midpoint of a spherical mirror, like the one shown in Concept 1, is called the *principal axis* of the mirror. This line is perpendicular to the surface of the mirror where they intersect.

While planar mirrors produce only virtual images, a curved mirror can produce either a virtual or a real image, depending on the shape of the mirror and the location of the object.

The *image point* corresponding to an object on the principal axis is the location on the axis of its image. When an object is infinitely far away from the mirror, its image point defines the *focal point* of a mirror or lens. You see this point, F, in Concept 2. "Infinitely far away" means the incident light rays from the object are effectively parallel. The focal point is located on the principal axis, and it is an unchanging characteristic of a particular mirror.

concept 1

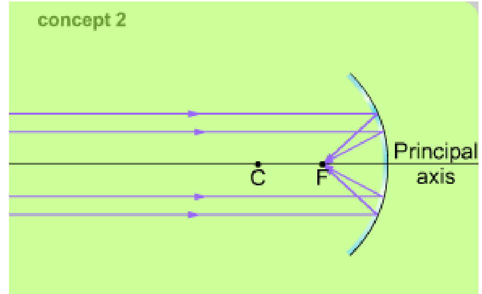


Mirror terminology
Center of curvature C is center of sphere
Principal axis passes through center of curvature and midpoint of mirror

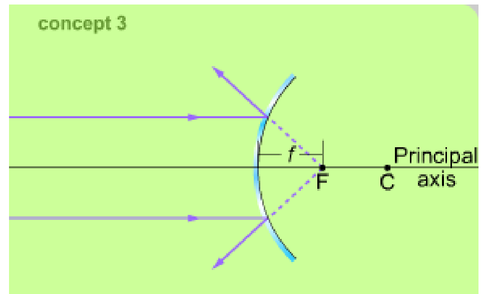
The distance between the focal point and the midpoint of a mirror is its *focal length*, represented with the symbol f. By convention, concave mirrors have positive focal lengths, while convex mirrors have negative ones. You see an example of a negative focal length in Concept 3. This convention allows the use of a single equation to locate images and objects for both types of mirrors.

Our study of mirrors will focus on *paraxial rays.* Paraxial rays are incident rays that are relatively close to the principal axis. Rays that are far from the principal axis and do not converge at a single image point are called *nonparaxial rays*. Nonparaxial rays cause blurry images. This effect is called *spherical aberration*.

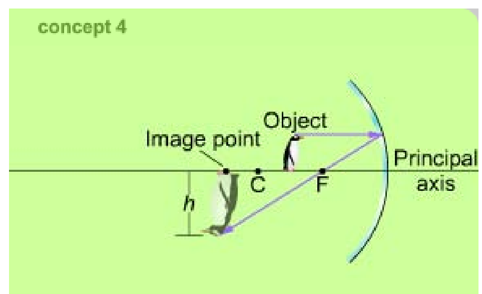
For our purposes, the base of an object is located on the principal axis. When an object's base is on the principal axis, then the base of its image will be too, though it may be inverted, like the image shown in Concept 4. The *height* of such an object or image is measured from the principal axis. Like most quantities associated with objects and images in optics, height has a sign. A positive value means the image is upright. Its top is still on top. A negative sign means an image is inverted, as you see in Concept 4. Its base still lies on the principal axis, but its top is now below the axis.



Focal point (F)
Image point of infinitely distant object



Focal length (*f*)
Distance from mirror to focal point
· Positive on object side
· Negative on far side



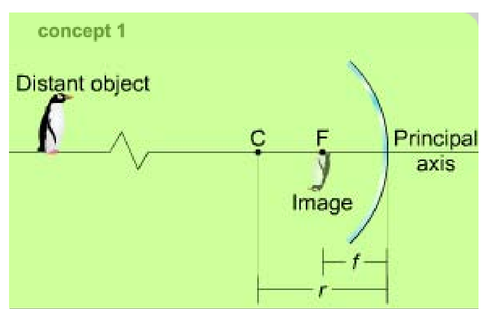
Height (*h*)
Image point and height *h* describe image

35.10 - Spherical mirrors: focal length equation

In Concept 1, you see a diagram of a concave mirror with center of curvature C and radius of curvature r. A distant object and its image are shown. (The object and image distances as well as the relative heights are not drawn to scale.) Because the object is far away, the image point is at the focal point. The focal length f is the distance from the mirror to the focal point.

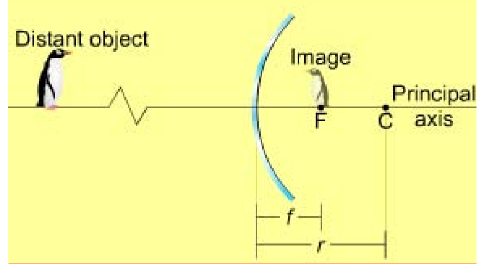
How does the focal length change as the mirror radius changes? Imagine you are using a mirror to create an image of a distant object. A concave mirror with a very slight curvature would correspond to a sphere with a large radius. If you consider the law of reflection and how rays would reflect from this mirror, you will realize that the focal point (and image) would be relatively far away from it. In contrast, a sharply curving mirror would create an image quite close to the surface.

The equation shown in Equation 1 quantifies this general relationship. The magnitude of the focal length f equals one-half the radius r of the sphere. It is positive for a concave mirror, and negative for a convex mirror.



Focal length
Positive for concave mirror
Negative for convex mirror

equation 1



Focal length

Concave spherical mirror:

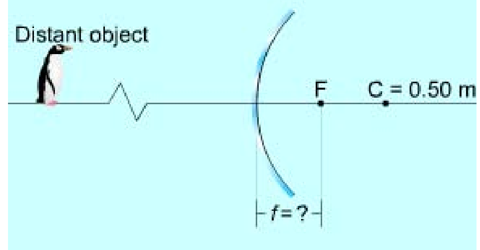
$$f = r/2$$

Convex spherical mirror:

$$f = -r/2$$

f = focal length
r = distance to center of curvature

example 1



What is the mirror's focal length?

$f = -r/2$

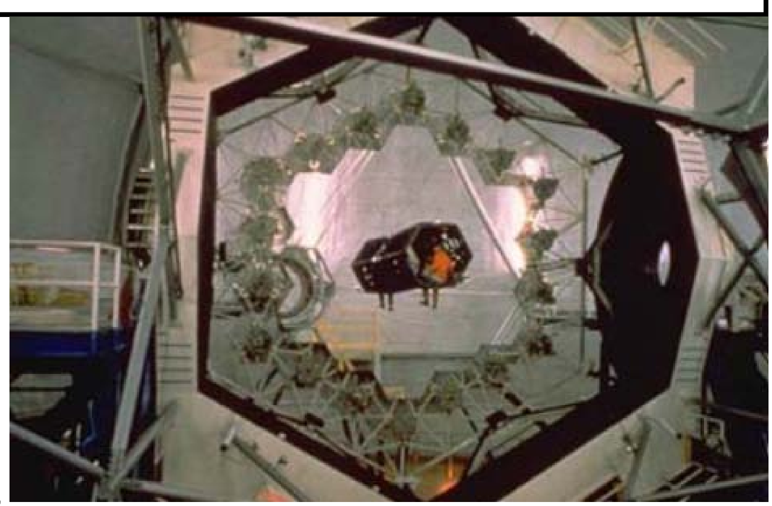
$f = -(0.50\text{ m})/2$

$f = -0.25\text{ m}$

35.11 - Parabolic mirrors

Parabolic mirrors provide a solution to the problem of spherical aberration. Spherical aberration occurs because some reflected rays do not pass precisely through the focal point of a spherical reflector. The farther the incident rays are from the mirror's principal axis, the more their reflected rays will miss the focal point. The result is a blurry image.

Spherical aberration is a serious problem for astronomers and others who need high quality images. Astronomers use telescopes to create images of distant objects. These telescopes often include very large concave mirrors to collect as much light as possible, allowing them to create images of objects that are exceedingly dim. (And by "large" we truly mean large: The diameter of a modern reflecting research telescope is in the eight to ten meter range.) If spherical mirrors were the basis of such telescopes, the images they created would be blurry due to spherical aberration.



This telescope mirror is a hexagonal mosaic of smaller spherical mirror segments. Note the workman sitting at the focal point.

To produce a sharp image, some telescopes instead use a mirror that has a cross section in the shape of a parabola rather than a section of a circle. Why? Reflected rays from distant objects converge to form a sharp image, no matter how far the incident ray is from the principal axis.

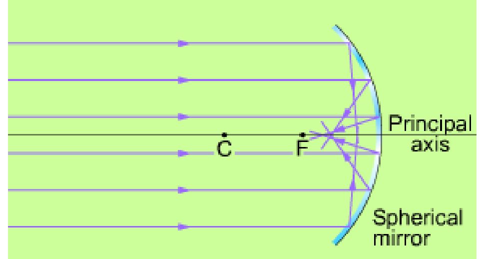
The parabolic design is used not just for visible light but for other forms of electromagnetic radiation, like radio waves, as well.

For instance, the *dish antennas* used for satellite television are small parabolic reflectors. These antennas do not have to be as smooth as light reflectors because the wavelength of the television signals they receive is far greater than the wavelengths of visible light. For electromagnetic radiation of even longer wavelengths, like the longest radio waves, the receiving parabolic reflectors can be even less smooth, but they must be very wide. You see an example of a large radio telescope in Concept 3.

Mirrors having a precise parabolic shape are difficult to manufacture. Some modern optical telescopes, called mosaic telescopes, combine the best features of both spherical and parabolic mirrors. They consist of a number of spherical mirror segments arranged in an overall parabolic shape. The mirrors are cut into hexagons so that they fit together neatly with no gaps. In the photograph above you see such a mirror.

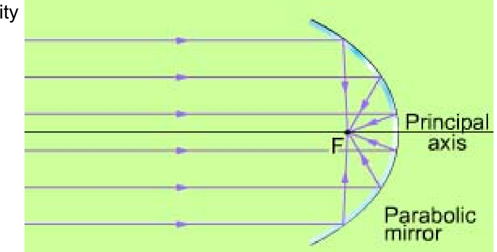
This arrangement works well because the individual spherical segments are small enough to have negligible aberrations of their own and to be easy to assemble, and like all spherical mirrors they are easy to manufacture. Yet each small mirror can be positioned as part of a large overall parabolic shape that optimizes the net image quality produced by the reflector.

concept 1



Spherical mirrors
Suffer from spherical aberration

concept 2

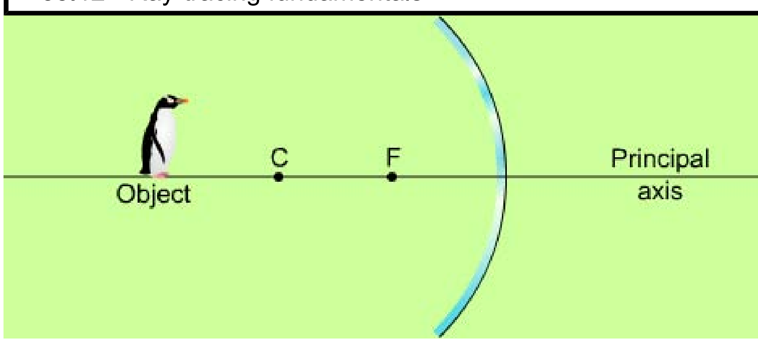


Parabolic mirrors
Do not exhibit spherical aberration

concept 3

Parabolic reflectors in use
Common in telescopes and satellite dishes

35.12 - Ray-tracing fundamentals



concept 1

What will be the nature of the image?

Ray tracing enables you to determine the fundamental properties of an image: Is it upright or inverted? Smaller or larger than the object? Real or virtual?

Ray tracing also provides an effective tool for verifying the computational results you get from applying the mirror and lens equations. It is easy to drop a sign, forget a reciprocal, or make other errors when applying these equations; ray tracing allows you to subject your answers to a “reality check.”

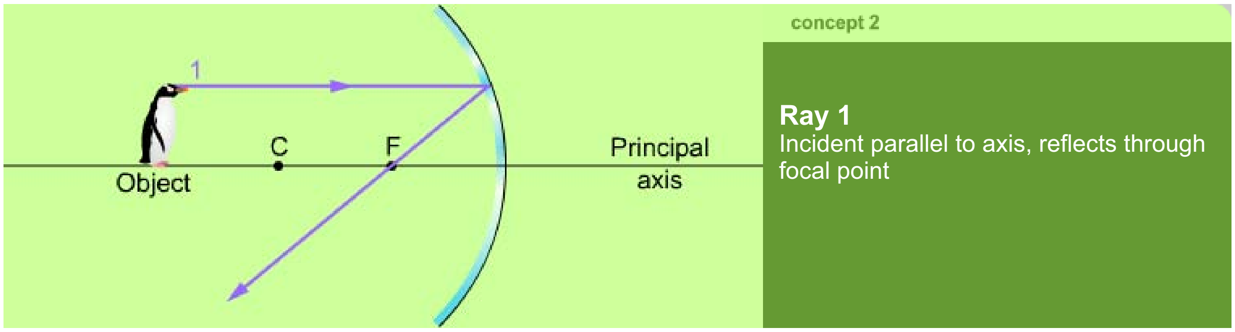
In this section, ray tracing will be used to determine the basic attributes of an image that is produced by an object located farther from a concave mirror than the center of curvature. You see this configuration above.

To locate the image, we start with three rays emanating from the top of the object: in this case, the top of the penguin’s head. The rays reach the mirror at three different points. They then reflect and converge. The point at which they converge is the location of the top of the image penguin's head.

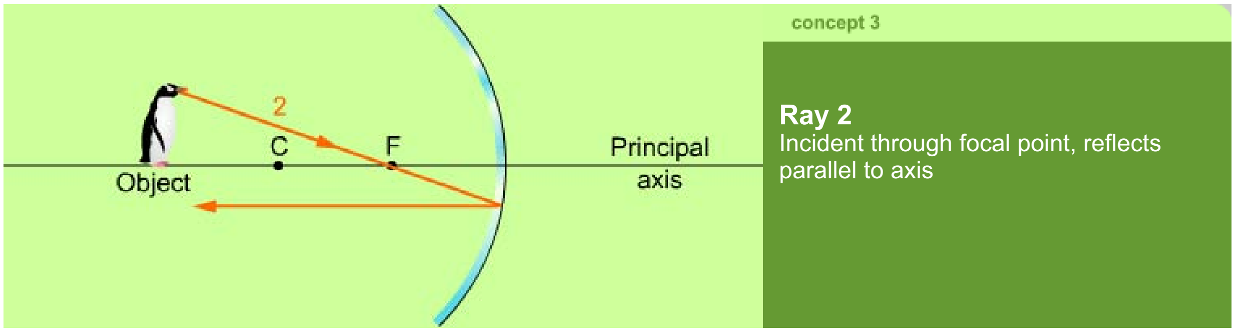
With hand-drawn ray-tracing diagrams, we use the working approximation that the incident rays are paraxial, so that they converge at one point. In the simulations in this chapter, you can observe accurate ray-tracing diagrams that display the effects of spherical aberration.

Below, we show each ray separately; all three rays and the image are shown in Concept 5.

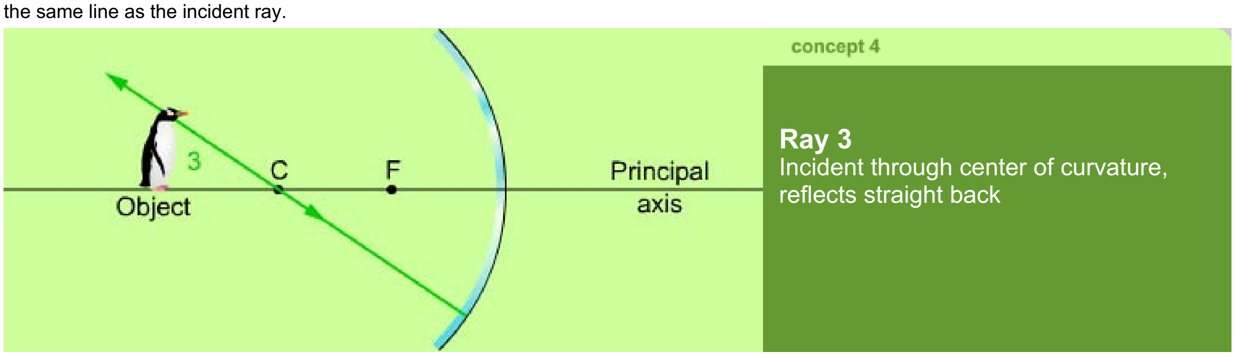
Ray 1 starts as an incident ray that is parallel to the principal axis. It reflects off the mirror and passes through the focal point after it reflects.



Ray 2 starts as an incident ray that passes through the focal point and then reflects parallel to the principal axis.



Ray 3 begins as an incident ray that passes through the center of curvature, strikes the mirror perpendicularly, and reflects back, moving along the same line as the incident ray.

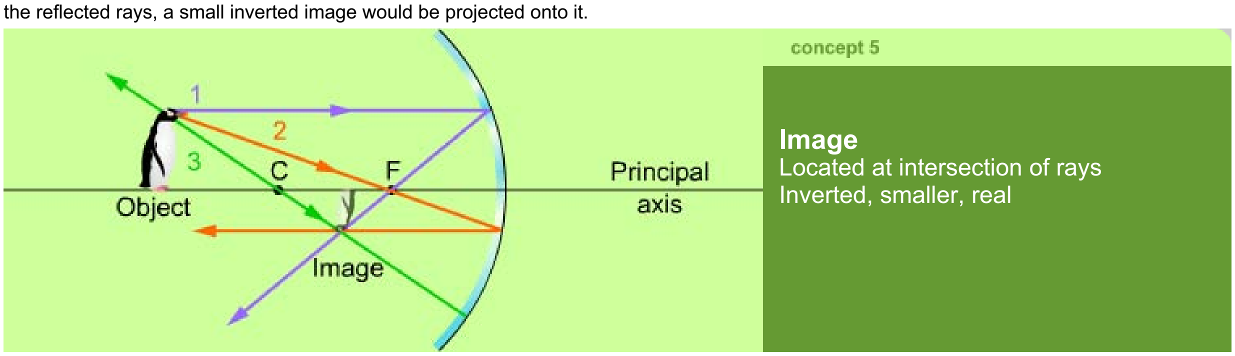


Why do these rays reflect in the way they do? The first incident ray is parallel to the principal axis. By definition, the reflection of this ray passes through the focal point.

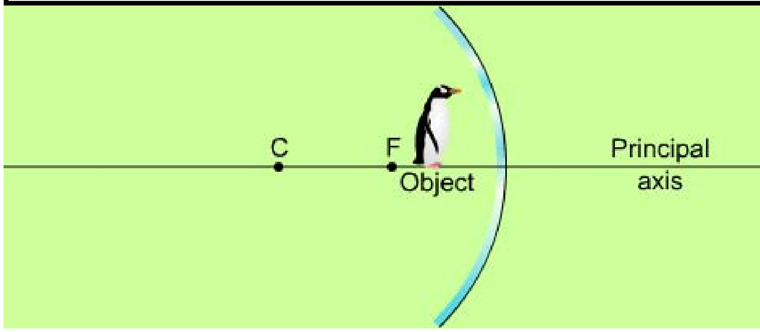
The second ray is parallel to the principal axis **after** it reflects. The definition of focal point applied "in reverse" explains why this should be so: Since the ray passes through the focal point before it reflects, it is parallel to the axis after it reflects.

The final ray reflects straight back because, like a radius of the sphere, it meets the spherical surface perpendicularly. Its angle of incidence is zero degrees, so its angle of reflection will be the same.

Now that the three rays have been drawn separately, we draw them together in Concept 5. They converge at the head of the image. The image is both inverted (upside down) and smaller than the original object. It is located between the focal point and the center of curvature. The image is real, being located on the same side of the mirror as the object. Since it is real, if a piece of paper were placed at the point of intersection of the reflected rays, a small inverted image would be projected onto it.



35.13 - Ray tracing: a second example



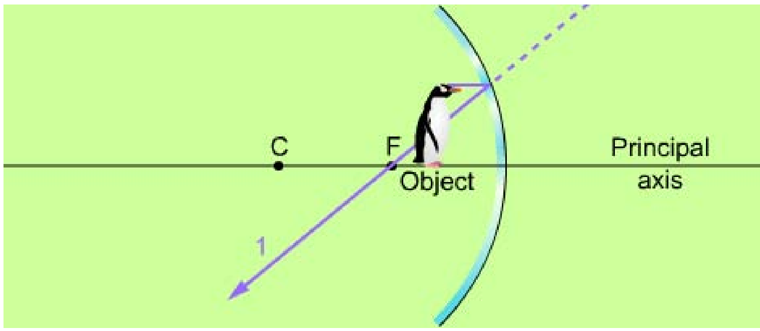
concept 1

What will be the nature of the image?

We will use ray tracing to determine the nature of the image created by an object that lies within the focal point of a concave mirror. As with the rays in the prior example, the incident rays all start or pass through a point at the top of the object, the penguin. They also have the same basic properties as the previous rays: One is parallel to the principal axis, one passes through the focal point, and the third passes through the center of curvature.

In contrast to the prior example, the reflected rays do not converge. We will need to use virtual rays to locate the image.

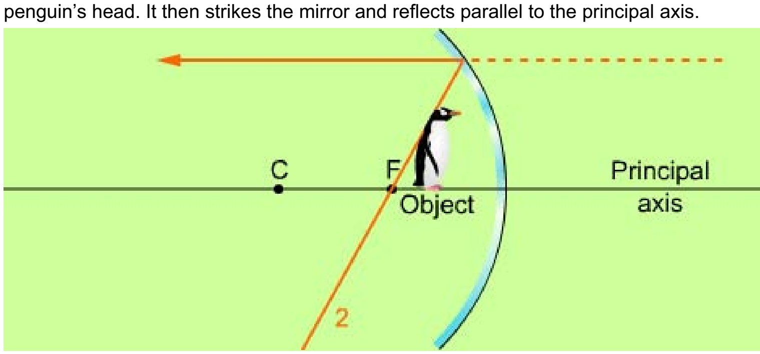
Ray 1 starts parallel to the principal axis, reflects, and passes through the focal point. Note that we extend the reflected ray backward through the mirror surface as a virtual ray (dashed line).



concept 2

Ray 1
Incident parallel to axis

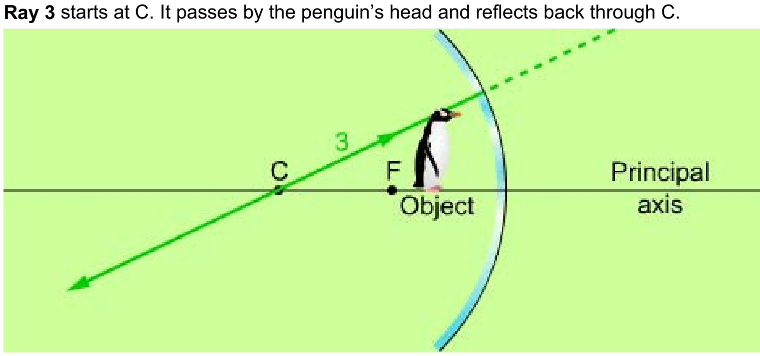
Ray 2 must pass through the focal point before reaching the mirror. We draw it as passing through the focal point before intersecting the penguin's head. It then strikes the mirror and reflects parallel to the principal axis.



concept 3

Ray 2
Passes through focal point

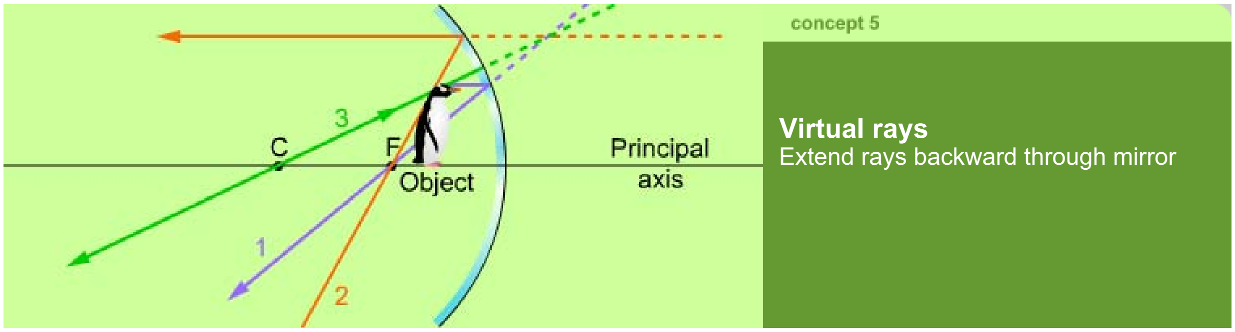
Ray 3 starts at C. It passes by the penguin's head and reflects back through C.



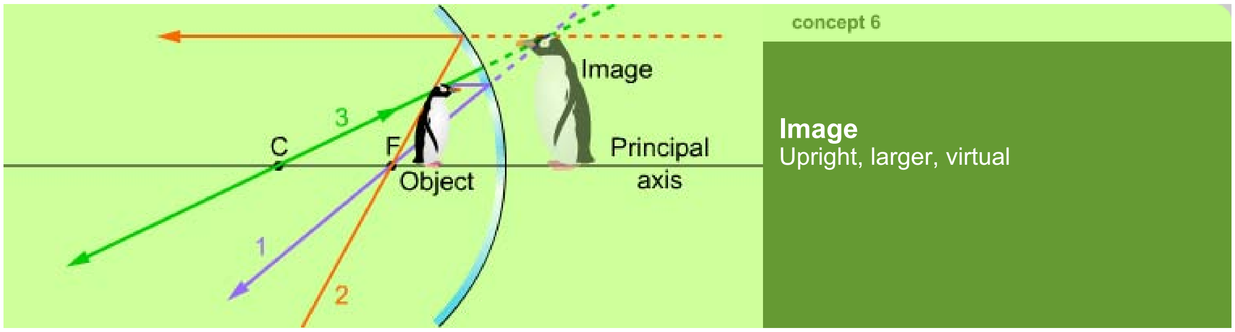
concept 4

Ray 3
Passes through center of curvature

In Concept 5, you see all three rays in the same diagram. The real rays do not converge; to locate a point of convergence we must use the virtual rays shown in the illustration.

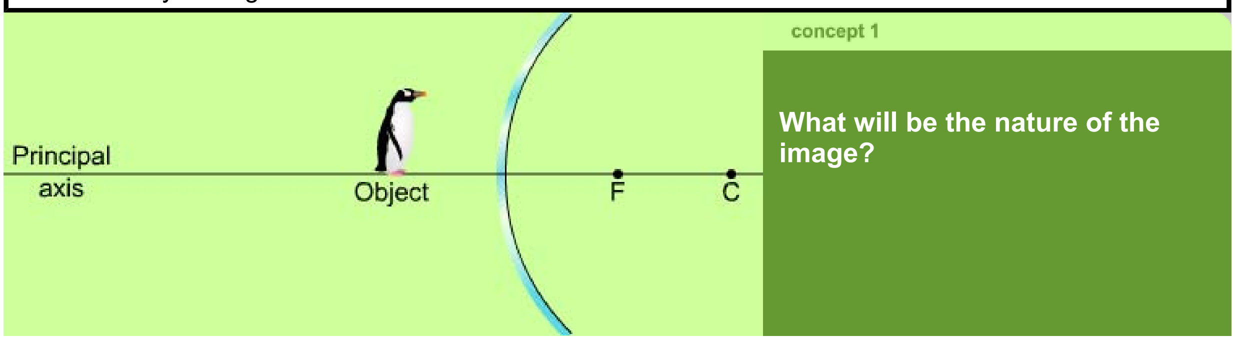


The point where the virtual rays converge, behind the mirror, is where the top of the image is located. We show the image in Concept 6. It is upright, larger than the object, and virtual.



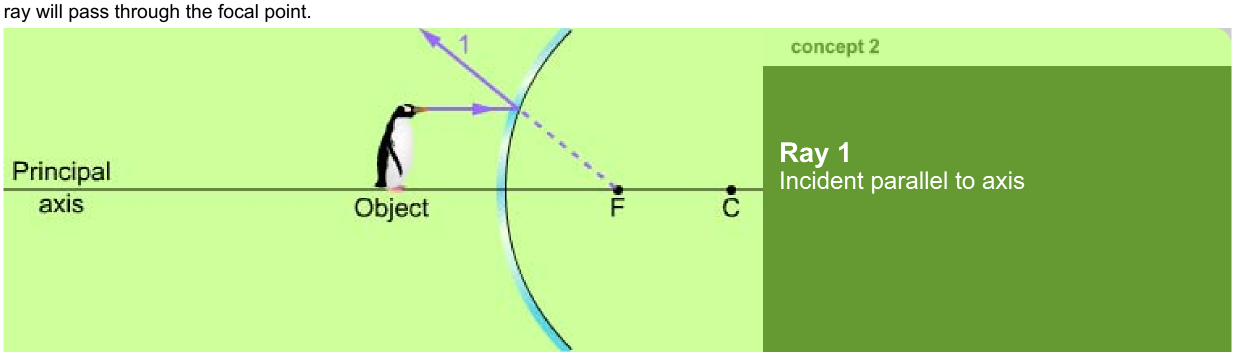
When you look at yourself in a concave makeup mirror, you put your face within its focal point to see a magnified image of yourself.

35.14 - Ray tracing: a convex mirror

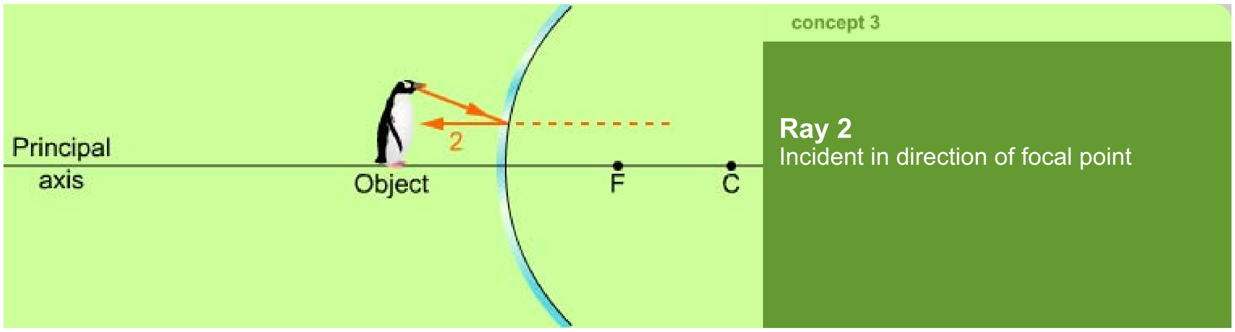


In this section, we construct a ray-tracing diagram for the image created by a convex mirror. The focal point and the center of curvature are always **behind** the reflecting surface of a convex mirror. We use three rays with the same essential properties as before, modifying their construction to compensate for the fact that the two points are on the opposite side of the mirror from the object.

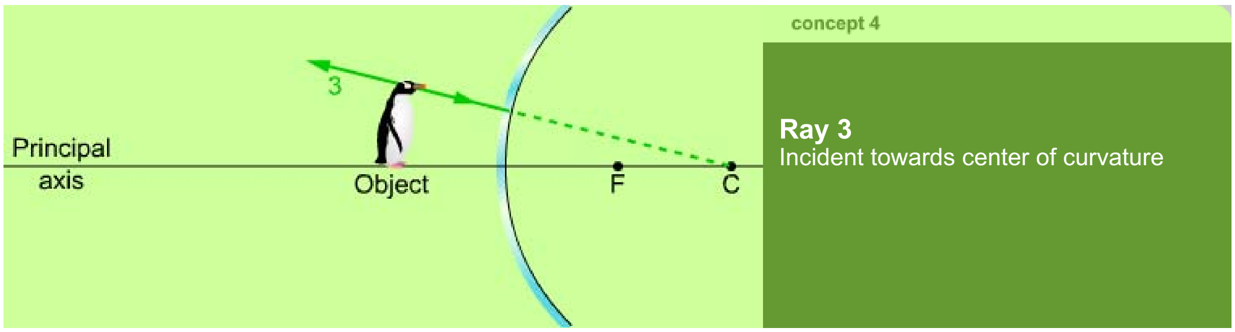
Ray 1. Ray 1 is incident parallel to the principal axis. If we extend the reflected component of this ray backward through the mirror, the virtual ray will pass through the focal point.



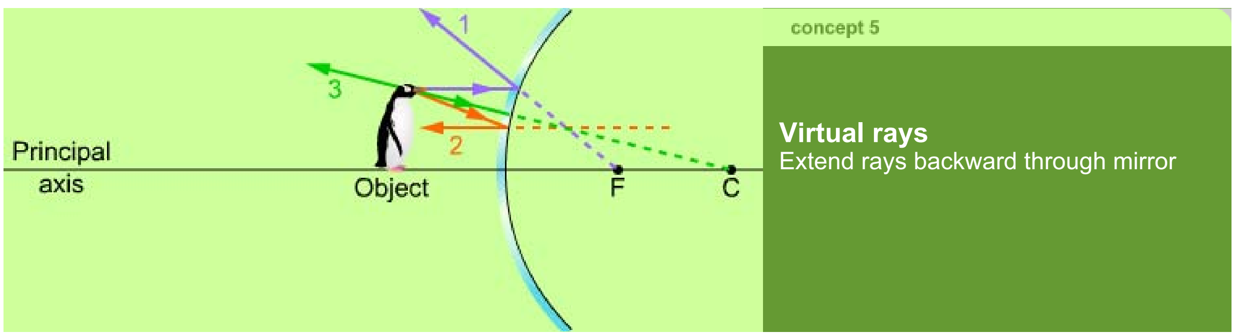
Ray 2. Instead of passing through the focal point, the incident part of ray 2 is directed toward it. Before it can reach the focal point behind the mirror, it reflects parallel to the principal axis. Its virtual extension behind the mirror is also parallel to the axis.



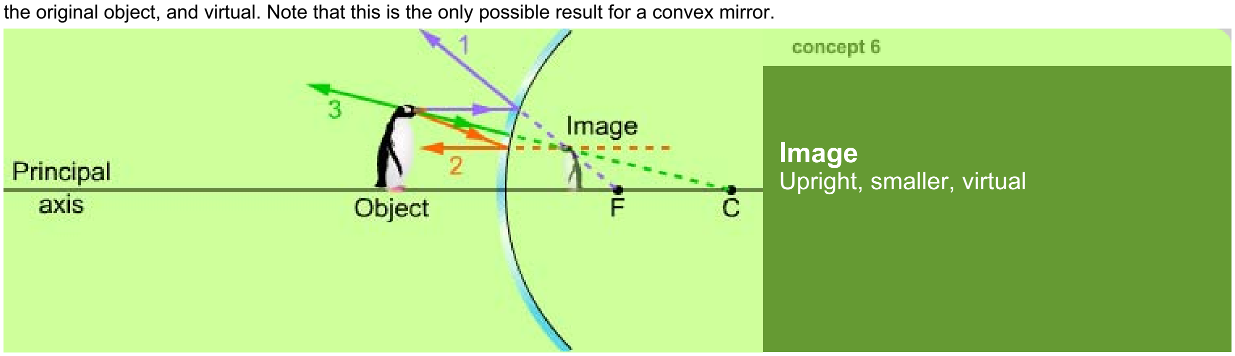
Ray 3. The incident component of Ray 3 is directed toward the center of curvature on the far side of the mirror and reflects back along the same line. The virtual extension of the reflected ray passes through the center of curvature.



In the following illustration we draw all three rays together. There is no intersection of the actual reflected rays in front of the mirror, but their backward virtual extensions intersect behind the mirror.



Finally, in Concept 6, we show the image as determined by the convergence point of the virtual rays. The image formed is upright, smaller than the original object, and virtual. Note that this is the only possible result for a convex mirror.



This ray-tracing diagram can be used to explain why security and automobile passenger-side mirrors are often convex. The mirror forms viewable images when the corresponding objects extend quite far above (or below) the principal axis, providing a very wide field of view. As you know, mirrors like this, typically used on the passenger side of vehicles, are often labeled with safety warnings that state, “Objects in mirror are closer than they appear.”

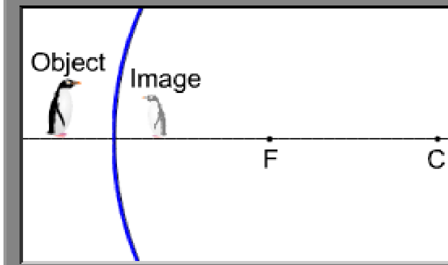
However, if you look at concept 6, you will note that the object is actually *farther* from the mirror than the virtual image is, and this is always true for images formed by a convex mirror. So are the warnings wrong? No, the convex mirror’s virtual image is always smaller than the image that would be produced by a planar mirror. Since the human brain relates size to distance, the image **appears** distant: We do interpret the object as being farther away than it really is.

35.15 - Interactive problem: image in a convex mirror

Here you use an interactive simulation to view the image produced by a convex mirror. You drag the object farther from and closer to the mirror and observe its image. You can turn on ray tracing with the SHOW RAYS button, and see how the rays can be used to determine the location of the image.

Some questions to consider include: Is the image ever inverted? Is it ever larger than the object? Is it ever real? Is it ever farther from the mirror than the object is? The object can be moved to positions that allow you to answer all these questions.

interactive 1



Convex mirror
See relationship of image and object

35.16 - Mirror equations

Quantity	Positive sign	Negative sign
Focal length, *f*	Concave mirror	Convex mirror
Image distance, d_i	In front of mirror (real)	Behind mirror (virtual)
Object distance, d_o	In front of mirror (real)	Behind mirror (virtual)
Magnification, *m* and height, *h*	Image upright	Image inverted

concept 1

Variables in the mirror equations
Interpretation of signs

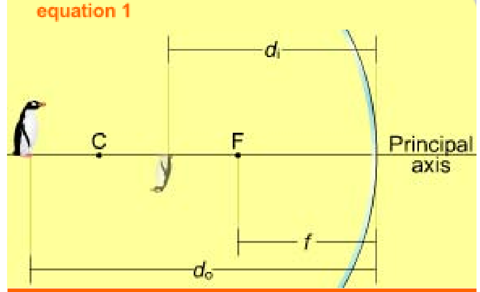
In this section, we discuss several equations used to determine the nature of images produced by mirrors. Before doing so, we will review and explain the mathematical sign conventions used in these equations.

Some of the notation we have already introduced: d is used for distance, f for focal length, and h for height. More specifically, d_o represents the distance of the object from the mirror and d_i the distance of the image from the mirror. Also, h_o is the height of the object and h_i is the height of the image.

With a single mirror, the object distance is always positive, although it can be negative in the more complex optical systems discussed below. The image distance is positive when the image is real, on the same side of the mirror as the object from which the image is created. The image distance is negative when the image is virtual, on the opposite side of the mirror. Remember that the focal length is positive for concave mirrors and negative for convex mirrors. The image height is positive for upright images and negative for inverted ones.

The magnification, represented by m, equals the image height divided by the object height. When the magnification is positive, the image is upright; when it is negative, the image is inverted relative to the object. Sometimes, the magnification as defined here is called *lateral magnification* to make it clear that it represents the change in height. When m is greater than one, then the image is larger than the object: what people ordinarily think of as "magnified."

equation 1



Mirror equation

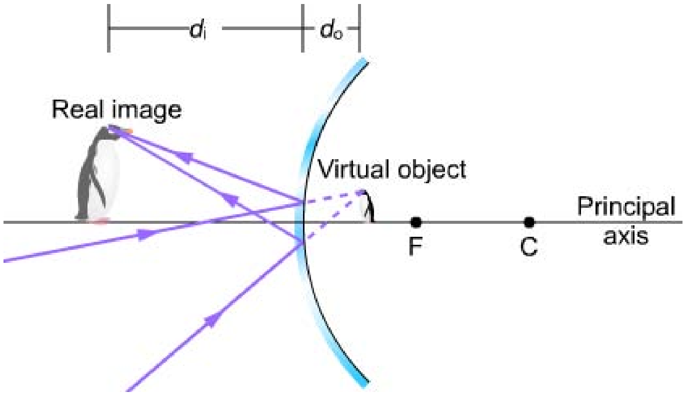
$$\frac{1}{d_i}+\frac{1}{d_o}=\frac{1}{f}$$

d_i = image distance
d_o = object distance
f = focal length

Now, on to the equations. The mirror equations are very useful for designing equipment such as cameras and telescopes because they enable engineers to correctly focus an image at a required point (for example, at the surface of a photographic film, or in a digital camera, at the surface of a "charge-coupled device"). The formula in Equation 1 shows the relationship of object and image distance to the focal length of a mirror. It is valid when the incident rays from the object are paraxial.

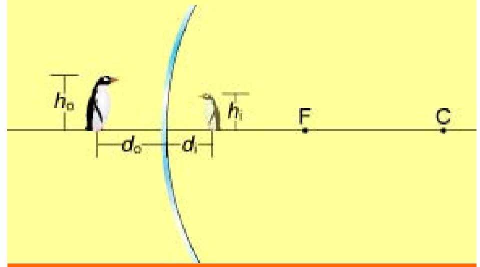
The next two equations, shown in Equation 2, are used to define and calculate magnification. The first equation defines magnification: It is the ratio of the image and object heights. The magnification can also be calculated as the negative of the ratio of image distance to object distance.

If you now refer back to the table in Concept 1 for signs, you may see one entry that surprises you: an object with a negative distance. Such a "virtual object" can be created by a configuration of two mirrors, or by a lens and a mirror, as shown in the illustration below.



Incident rays from the original real object first reflect off some mirror, or pass through some lens, which is not shown in the illustration, being off screen to the left. As a result, converging rays come in from the left and strike the convex mirror in the diagram. The image they create would be a real image "behind" the location of the mirror if the mirror were not there. Since the mirror is there, the convergence point of the virtual rays on the right defines the location of a virtual object, whose distance from the mirror is negative. The incident rays reflect off the front of the mirror to create a real image, with a positive image distance, as shown. Since optical instruments such as telescopes and microscopes often rely on a combination of lenses and mirrors, negative object distances are not rare.

equation 2



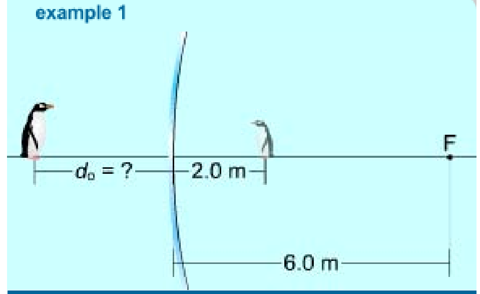
Magnification equations

$$m = \frac{h_i}{h_o}$$

$$m = -\frac{d_i}{d_o}$$

m = magnification of image
h_i = height of image
h_o = height of object

example 1



What is the distance to the object?

$d_i = -2.0$ m, $f = -6.0$ m

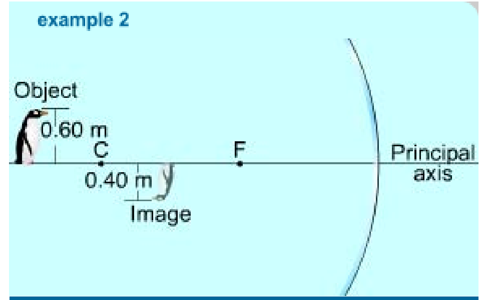
$$\frac{1}{d_i} + \frac{1}{d_o} = \frac{1}{f}$$

$$\frac{1}{d_o} = \frac{1}{f} - \frac{1}{d_i}$$

$$\frac{1}{d_o} = \frac{1}{-6.0\ \text{m}} - \frac{1}{-2.0\ \text{m}}$$

$$\frac{1}{d_o} = 0.33\ \text{m}^{-1}$$

$d_o = 3.0$ m



What is the magnification?

$h_o = +0.60$ m, $h_i = -0.40$ m

$m = h_i / h_o$

$m = -0.40 \text{ m} / 0.60 \text{ m}$

$m = -0.67$

35.17 - Interactive checkpoint: mirror equations

A frog sits in front of a convex spherical mirror. The mirror produces an image of the frog with a lateral magnification of 1/4. The mirror's radius of curvature is 0.260 m. How far away is the frog from the mirror?

Answer:

$d_o =$ ______ m

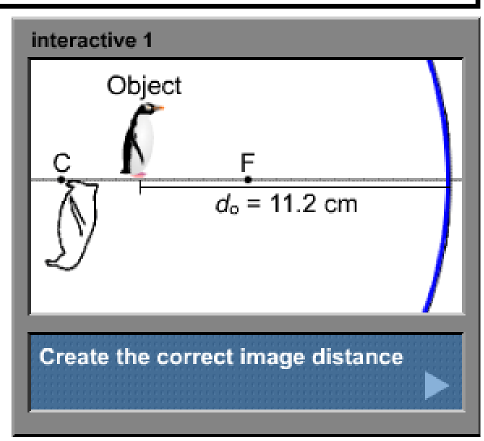
35.18 - Interactive problem: optical bench with a mirror

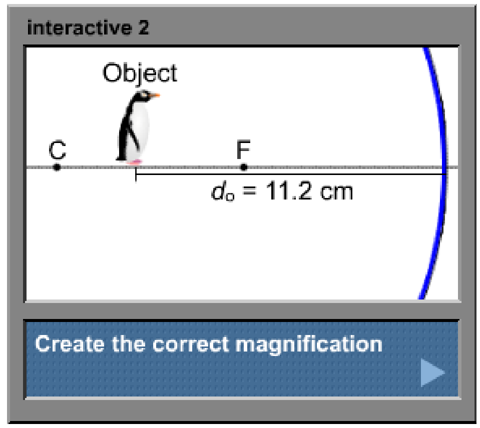
The simulations here challenge you to use the mirror equations. In the first, you are asked to create an image 13.7 cm from the mirror's surface as shown in the illustration and in the simulation. The object is 11.2 cm from the mirror.

You set the focal length of the mirror by dragging the focal point F, or by setting its value in the control panel. As you change the focal length, you will also be changing the curvature of the mirror. Once you believe you have set the focal length correctly, press the CHECK button to test your answer. Be careful with signs!

Your mission in the second simulation to the right is to create a mirror that will have a magnification of +0.470 for the same object at the same position. Again, you control the focal length, and to test your answer, press the CHECK button in this simulation. As a hint: First, determine where the image must be using one formula, then use another formula to determine the required focal length.

If you have trouble with either of these tasks, review the prior section on the mirror equations.





35.19 - Gotchas

Virtual images do not exist. This may be a philosophical question. They do exist in our minds. They cannot be touched, or projected onto a piece of paper. In contrast, real images can be projected onto a screen or a piece of paper.

A convex and a concave mirror have the same radius of curvature. *This means their focal lengths are identical.* The magnitudes of their focal lengths are the same, but the signs differ. The focal length is positive for a concave mirror and negative for a convex mirror.

A virtual image has a negative image distance. Yes. Conversely, real images have positive image distances.

35.20 - Summary

Reflection occurs when light bounces back from a surface, rather than passing through it or being absorbed. Reflection causes objects to be visible.

In optics, it is often useful to represent the path of light in the form of a light ray, a straight line that can be used to determine how light is affected by mirrors and lenses.

Planar mirrors are the simplest and most common type of mirror. In a planar mirror, your image has the same dimensions as you do and it appears to be the same distance behind the mirror as you are in front of it. The image is reversed front to back, but not right to left or top to bottom.

When you see your image in a planar mirror, you are seeing a virtual image, so called because you cannot project this image onto a screen at its perceived location. A virtual image appears to be on the opposite side of a mirror from the object it is created from, and its image distance is negative.

In contrast, a real image **can** be projected onto a screen, and it occurs on the same side of a mirror as the object it is created from. Its image distance is positive. Curved mirrors can create virtual or real images.

The law of reflection states that for a light ray reflecting off a mirror, the angle of incidence equals the angle of reflection. These two angles are measured between the light ray and a line normal to the surface of the mirror.

A ray diagram can help you determine the location of the image an object casts in a **planar** mirror. Draw two incident rays from any point on the object to the mirror. Draw the reflected rays using the law of reflection, and trace their virtual extensions backward behind the mirror as dashed lines. The point of the image corresponding to the point source of the incident rays on the object is located where these virtual rays intersect.

Equations

Law of reflection

$\theta_{\mathrm{i}} = \theta_{\mathrm{r}}$

Focal length of concave mirror

$f = r/2$

Focal length of convex mirror

$f = -r/2$

Mirror equation

$$\frac{1}{d_{\mathrm{i}}} + \frac{1}{d_{\mathrm{o}}} = \frac{1}{f}$$

Magnification equations

$$m = \frac{h_{\mathrm{i}}}{h_{\mathrm{o}}} = -\frac{d_{\mathrm{i}}}{d_{\mathrm{o}}}$$

Spherical mirrors have a constant curvature. A mirror is called concave if it curves toward (around) the object it reflects. A convex mirror curves away from the object it reflects.

All points on a spherical mirror are equidistant from its center of curvature. The principal axis is a line through the center of curvature and the midpoint of the mirror. Paraxial rays are incident rays close to the principal axis of a mirror. The reflections of these rays converge at the focal point if the rays are parallel to the principal axis. The height of an image is negative if the image is inverted.

The focal length of a curved mirror is the distance between a mirror and its focal point. By convention, it is positive for a concave mirror where the focal point is on the same side of the mirror as the object. It is negative for a convex mirror where the focal point is on the opposite side.

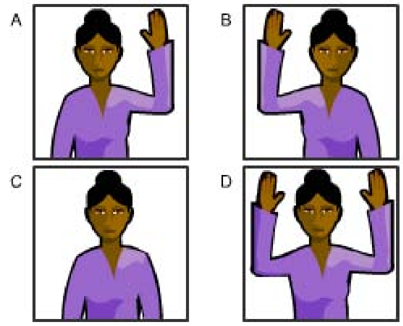
Spherical mirrors exhibit spherical aberration because only paraxial incident rays converge at the focal point. With a parabolic mirror, **all** incident rays that are parallel to the principal axis converge at the focal point.

Ray tracing for concave and convex mirrors is more complicated than for planar mirrors, but it can help you determine the orientation, relative size, and location (or type) of an image.

The mirror equation relates the distances of an object and its image from a mirror to the focal length of the mirror. The magnification equations define the magnification of an image created by a mirror. If the magnification is known, we may calculate the image height from the object height or the image distance from the object distance, or vice versa.

Conceptual Problems

C.1 You are standing in a cubical room where two adjacent walls are mirrors. You stand so that you are equidistant from the two mirrored walls. (a) How many images of yourself can you see? (b) You raise your right hand and face the corner where the mirrors meet. Which of the following sketches best represents what you will see?



(a) i. 0
ii. 1
iii. 2
iv. 3

(b) i. A
ii. B
iii. C
iv. D

C.2 An object is placed between the focal point and center of curvature of a concave mirror. Draw a ray diagram to determine the nature of the image. Check all the boxes that describe the image.

- ☐ Real
- ☐ Virtual
- ☐ Smaller
- ☐ Larger
- ☐ Upright
- ☐ Inverted

C.3 An object is placed between a concave mirror and its focal point. Draw a ray diagram to determine the nature of the image. Check all the boxes that describe the image.

- ☐ Real
- ☐ Virtual
- ☐ Smaller
- ☐ Larger
- ☐ Upright
- ☐ Inverted

C.4 An object is placed in front of a convex mirror. Draw a ray diagram to determine the nature of the image. Check all the boxes that describe the image.

- ☐ Real
- ☐ Virtual
- ☐ Smaller
- ☐ Larger
- ☐ Upright
- ☐ Inverted

C.5 An object is placed 20.0 cm in front of a concave mirror with focal length 9.00 cm. Draw a ray diagram to determine the nature of the image. Check all the boxes that describe the image.

- ☐ Real
- ☐ Virtual
- ☐ Smaller
- ☐ Larger
- ☐ Upright
- ☐ Inverted

C.6 If you want to create a real image, which object location and mirror combinations could you use? Check all that apply.

- ☐ Object in front of convex mirror
- ☐ Object at distance over twice the focal length from concave mirror
- ☐ Object between focal point and center of curvature of concave mirror
- ☐ Object at distance less than focal length from concave mirror

C.7 You hold up your index finger very close to the surface of a concave mirror (inside the focal point). (a) As you move your finger farther from the mirror, does the image of your finger get larger or smaller? (b) What happens to the image when the distance between your finger and the mirror exceeds the focal length of the mirror? Select all that apply.

(a) ○ Larger ○ Smaller

(b) ☐ Stays the same
☐ Inverts
☐ Becomes virtual
☐ Becomes real

C.8 Explain how two parallel facing mirrors in a carnival's house of mirrors create an infinite number of images.

C.9 As you are driving your car, you look in the passenger side mirror and see a car behind you. At the bottom of the mirror is a note from the manufacturer stating, "Objects in mirror are closer than they appear." What is the curvature of the mirror? Explain why objects appear farther away than they are in this type of mirror.

i. Planar
ii. Concave
iii. Convex

C.10 In a carnival House of Mirrors, you come face to face with an image of yourself that is right side up and twice your normal height. (a) What is the shape of this mirror? (b) Are you inside (closer to the mirror than) or outside (farther from the mirror than) the focal point?

(a) i. Planar
ii. Concave
iii. Convex

(b) i. Inside
ii. Outside

C.11 Planar mirrors do not treat left and right any differently than they do up and down. For example, the image of your right hand is directly across from the original, just as the image of your head and the image of your feet are directly across from the original objects. Mirrors do reverse front and back. Why, then, does "mirror writing" appear to be right-left reversed in a planar mirror?

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Where can you place the object so that the image and object are on the same side of the mirror? (b) Where can you place the object so that the image is smaller than the object? (c) Where can you place the object so that the image is larger than the object? Check all that apply.

(a) ☐ To the left of C
☐ In between C and F
☐ To the right of F

(b) ☐ To the left of C
☐ In between C and F
☐ To the right of F

(c) ☐ To the left of C
☐ In between C and F
☐ To the right of F

Section 3 - Mirror basics

3.1 As you get ready for your early morning physics class, you comb your hair in front of a planar mirror. You know that you are 3.75 m in front of the mirror. How far away from you is your image?

__________ m

3.2 Two parallel planar mirrors face each other. An object is placed between them at a distance *x* from one and 3*x* from the other. If the images created of the object are "first-order" images, and those created of the first-order images are "second-order" images, how far away are the second-order images from the object, in terms of *x*?

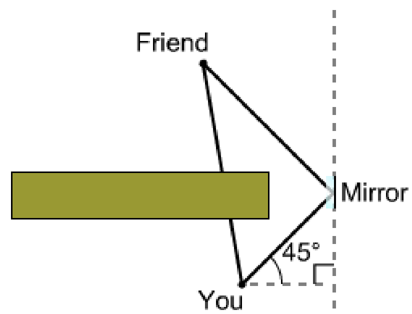
__________ x

Section 5 - The law of reflection

5.1 Two mirrors are placed at a 90° angle to each other. A light ray strikes one mirror 0.710 m from the intersection of the mirrors with an incident angle of 36.5°. The ray then travels a distance *d* before reflecting from the second mirror. (a) What is the distance *d*? (b) What is the angle with which the light ray is reflected from the second mirror? All angles are measured from the normal belonging to the appropriate mirror. (c) What is the angle between the direction of the incident ray and that of the final reflected ray? Express your answer as a positive number between 0° and 360°.

(a) __________ m
(b) __________ °
(c) __________ °

5.2 You and a friend are playing a game of squirt-gun tag in a maze. Suddenly you see your friend's image in a small planar mirror (see illustration). You take a shot over the barrier in front of you and find that your friend is just at the end of the 7.0 m range of your squirt gun. If you are 4.0 m from the point of reflection of the light ray, how far from the point of reflection of the light ray is your friend?



__________ m

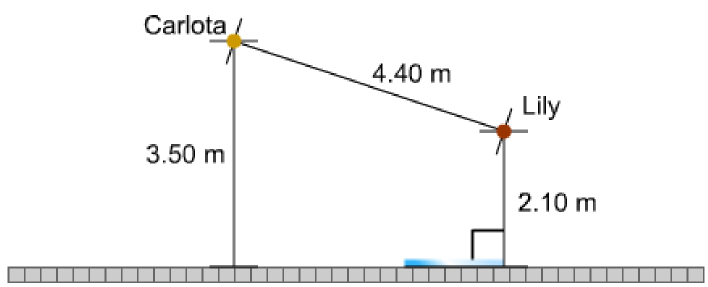
5.3 Two 3.0 meter-tall planar mirrors are placed facing and parallel to each other 0.75 m apart. If a ray of light just passes the bottom edge of one mirror and strikes the other at an incident angle of 10°, how many reflections will the light make before exiting this "hall"?

____

5.4 A bartender keeps an eye on her customers while her back is turned by way of a mirror mounted behind the bar. The mirror is 0.580 m wide and she is standing 0.750 m from it. How wide an expanse of the parallel wall opposite the mirror can she see in the mirror if the opposite wall is 6.00 m away from the mirror?

__________ m

5.5 Carlota is standing 3.50 m from a wall. Her friend Lily is standing 2.10 m from the same wall, and they are 4.40 m from each other. If a mirror is mounted on the wall with its edge directly in front of Lily and extending toward Carlota as in the illustration, how wide must the mirror be for the friends to see each other in the mirror?



__________ m

Section 7 - How tall should a full-length mirror be?

7.1 An emperor penguin is standing 2.75 m from a planar mirror. If it is 1.20 m tall with its eyes 6.00 cm below the top of its head, (a) what minimum height mirror will it need to see its entire reflection and (b) how far up from the floor should the bottom edge of the mirror be located?

(a) __________ m
(b) __________ m

Section 10 - Spherical mirrors: focal length equation

10.1 A shiny metallic Christmas ball ornament 4.60 cm in diameter is used as a mirror. What is the mirror's focal length?

__________ cm

10.2 You want to create a spherical mirror with a focal length of +15.0 cm. (a) Will the mirror be concave or convex? (b) What must be the radius of curvature?

(a) ○ Concave ○ Convex
(b) __________ cm

10.3 You are lost in the wilderness, carrying only a piece of aluminum foil. Thanks to diligent study in your biology class, you remember which mushrooms are poisonous, and you are able to pick others that won't kill you. Your talents in basket weaving also pay off magnificently, as you create a basket and devise a plan to line it with aluminum foil, making a solar cooker. Mmmm, roast mushrooms! The basket is a section of a sphere with radius 21.0 cm. How far from the bottom of the basket should you place the mushrooms to cook them most efficiently using the Sun's rays?

__________ cm

Section 15 - Interactive problem: image in a convex mirror

15.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Is the image ever inverted? (b) Is the image ever larger than the object? (c) Is the image ever real? (d) Is the image ever farther from the mirror than the object is?

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c) ○ Yes ○ No
(d) ○ Yes ○ No

Section 16 - Mirror equations

16.1 An object is 17.3 cm from a concave mirror of focal length 8.10 cm. (a) What is the image distance? State your answer with the correct sign. (b) What is the magnification? (c) Is the image upright or inverted?

(a) __________ cm
(b) __________
(c) i. Upright
ii. Inverted

16.2 Use the mirror equation to show that for a planar mirror, the magnitude of the image distance equals the magnitude of the object distance. Hint: A planar mirror can be thought of as a spherical mirror with an infinite radius of curvature.

16.3 The focal point of a convex mirror is 22.4 cm from the mirror's surface. An image seen in the mirror is between the mirror and the focal point, 12.9 cm from the focal point. How far from the mirror is the object?

__________ cm

16.4 A mirror creates an image of magnification −1/3. (a) Is the image upright or inverted? (b) If the object is 7.80 cm tall, what is the height of the image? State your answer with the correct sign.

(a) ○ Upright ○ Inverted
(b) __________ cm

16.5 A mirror produces an inverted image 2.80 cm tall when a 5.60 cm tall object is 25.0 cm away from it. (a) What is the image distance? State your answer with the correct sign. (b) What is the focal length of the mirror? State your answer with the correct sign.

(a) __________ cm
(b) __________ cm

16.6 You wish to use a concave mirror to produce a virtual image that is three times the size of an object. (a) Which will be closer to the mirror, the image or the object? (b) If the object is 22.0 cm from the mirror at this magnification, what is the radius of curvature of the mirror?

(a) i. Image
ii. Object
iii. They are equidistant from the mirror
(b) __________ cm

16.7 A planar mirror is placed at the origin on a number line with the reflective side facing the positive direction. A convex mirror with focal length −22.0 cm is placed at x = 35.0 cm with its reflective side facing the planar mirror. An object is placed at x = 19.0 cm. What is the distance between the image of the object created by the planar mirror and the image of the object created by the convex mirror?

__________ cm

16.8 A mirror of focal length 22.6 cm creates an image with magnification −0.350. What is the image distance?

__________ cm

16.9 An object and its image are both 12.4 cm from a spherical mirror, and on the same side of the mirror. (a) Is the mirror concave or convex? (b) What is the mirror's focal length? State your answer with the correct sign.

(a) ◯ Concave ◯ Convex
(b) __________ cm

16.10 You look into a shaving mirror and the upright image of your face is 40.0 cm away from you. The image has a magnification of 1.50. (a) How far away are you from the mirror? (b) What is the radius of curvature?

(a) __________ cm
(b) __________ cm

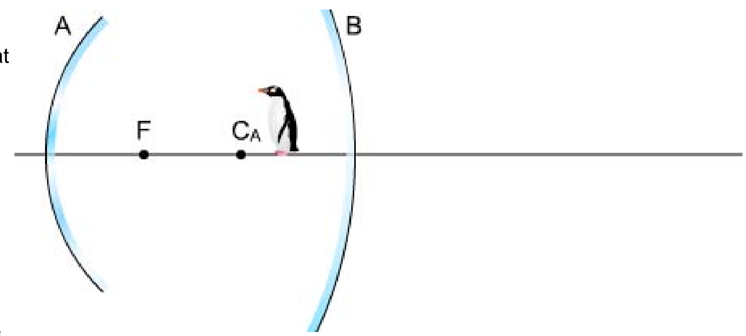
16.11 You place an object 34.0 cm in front of a convex mirror and the image produced is half the height of the object. What is the focal length of the mirror? State your answer with the correct sign.

__________ cm

16.12 A convex mirror has a radius of curvature of 36.0 cm. When a 5.60 cm tall object is placed in front of the mirror, the magnification is 0.250. (a) What is the height of the image? (b) How far is the object from the mirror?

(a) __________ cm
(b) __________ cm

16.13 Two concave mirrors are arranged as shown in the illustration. Their focal points coincide at point F, and the center of curvature of mirror A is indicated. An object is placed between the mirrors as shown. Consider the image created by mirror B of an image created by mirror A of the object. In other words, first imagine that mirror A creates an image of the object. The image created by A is then used as the object for mirror B to create the "final" image. Ignore all other images created by the mirrors. (a) Is the final image upright or inverted compared to the object? It may help to draw a ray diagram. (b) Is the final image real or virtual? (c) If the focal length of mirror A is 6.00 cm, the focal length of mirror B is 13.0 cm, and the object is 14.0 cm from mirror A, how far is the final image from the object (give a positive answer)?



(a) ◯ Upright ◯ Inverted
(b) ◯ Real ◯ Virtual
(c) __________ cm

16.14 An object is placed at position (2.00, 2.00) on a coordinate system. A concave mirror is placed at (−7.00, 2.00), facing the object. The mirror has a focal length of 2.25. At what position is the image produced? Assume that all lengths are in meters.

(__________ , __________)

Section 18 - Interactive problem: optical bench with a mirror

18.1 Use the information given in the first interactive problem in this section to calculate the focal length required for the image to appear at the desired location. Test your answer using the simulation.

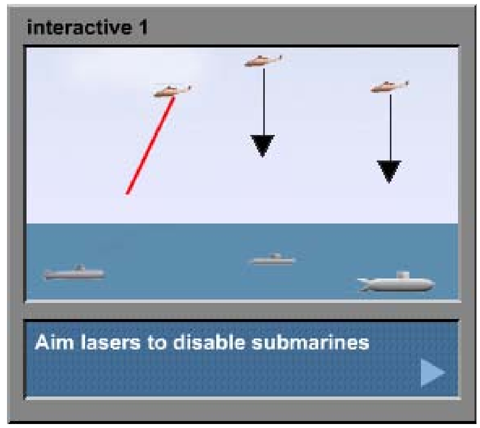
__________ cm

18.2 Use the information given in the second interactive problem in this section to calculate the focal length required to produce the desired magnification. Test your answer using the simulation.

__________ cm

chapter 36 Refraction

36.0 - Introduction



Light can refract − change direction − as it moves from one medium to another. For instance, if you stand at the edge of a pool and try to poke something underwater with a stick, you may misjudge the object's location. This is because the light from the object changes direction as it passes from the water to the air. You perceive the object to be closer to the surface than it actually is because you subconsciously assume that light travels in a straight line.

Although refraction can cause errors like this, it can also serve many useful purposes. Optical microscopes, eyeglass lenses, and indeed the lenses in your eyes all use refraction to bend and focus light, forming images and causing objects to appear a different size or crisper than they otherwise would. Where a lens focuses light, and whether it magnifies an object, is determined by both the curvature of the lens and the material of which it is made. Scientists have developed quantitative tools to determine the nature of the images created by a lens. We will explore these tools thoroughly later, "focusing" first, so to speak, on the principle of refraction underlying them.

To begin your study of refraction, try the simulation to the right. Each of your helicopters can fire a laser − a sharp beam of light − at any of three submarines lurking under the sea. The submarines have lasers, too, and will shoot back at your craft. Your mission is to disable the submarines before they disarm your helicopters. When you make a hit, you can shoot again. Otherwise, the submarines get their turn to shoot until they miss.

You play by dragging the aiming arrow underneath any one of your helicopters. Press FIRE and the laser beam will follow the direction of this arrow until it reaches the water, where refraction will cause the beam to change direction.

In addition to hitting the submarines before they get you, you can conduct some basic experiments concerning the nature of refraction. As with reflection, the angle of incidence is measured from a line normal (perpendicular) to a surface. In this case, the surface is the horizontal boundary between the water and the air. Observe how the light bends at the boundary when you shoot straight down, at a zero angle of incidence, or grazing the water, at a large angle of incidence. You can create a large angle of incidence by having the far right helicopter, for example, aim at the submarine on the far left.

You can also observe how refraction differs when a laser beam passes from air to water (your lasers) and from water to air (the submarines' lasers). Observe the dashed normal line at each crossover point and answer the following question: Does the laser beam bend toward or away from that line as it changes media? You should notice that the laser beams of the submarines behave differently than those of the helicopters when they change media.

As a final aside: You may see that some of the laser beams of the submarines never leave the water, but reflect back from the surface between the water and the air. This is called total internal reflection.

36.1 - Refraction

Refraction: The change in the direction of light as it passes from one medium to another.

Refraction bends water waves into line with the shore.

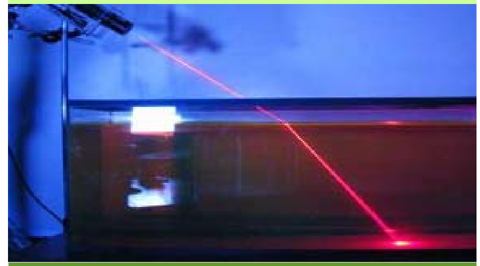
A material through which light travels is called a *medium* (plural: *media*). When light traveling in one medium encounters another medium, its direction can change. It can reflect back, as it would with a mirror. It can also pass into the second medium and change direction. This phenomenon, called refraction, is shown to the right. In the photo, a beam of light from a laser refracts (bends) as it passes from the air into the water.

Light refracts when its speeds in the two media are different. Light travels faster through air than in water, and it changes direction as it moves from air into water, or from water into air.

Although we are primarily interested in the refraction of light, all waves, including water waves, refract. Above, you see a photograph of surf wave fronts advancing parallel to a beach. Deep-ocean swells may approach a coastline from any angle, but they slow down as they encounter the shallows near the shore. The parts of a wave that encounter the shallow water earliest slow down first, and this causes the wave to refract. Sound waves can also refract. During a medical ultrasound scan, an acoustic lens can be used to focus the sound waves. The lens is made of a material in which sound travels faster than in water or body tissues.

The surface between two media, such as air and water, is called an *interface*. As with mirrors, light rays are often used to depict how light refracts when it meets an interface. Lasers are often used to demonstrate refraction because they can create thin beams of light that do not spread out. These light beams are physical analogues of light rays.

concept 1



Refraction
Change in wave direction at interface
Caused by change in speed of wave

36.2 - Index of refraction

Index of refraction of a material: The speed of light in a vacuum divided by the speed of light in the material.

Light travels at different speeds in different materials. The index of refraction of a material provides a measure of the speed of light in that material. The symbol n represents the index of refraction. The equation used to calculate the index of refraction for a material is shown to the right. The table to the right shows the indices for some common materials.

In the equation for the index of refraction, the speed of light in a vacuum is in the numerator. Light travels at its maximum speed in a vacuum, so the index of refraction for other materials is always greater than one.

The index of refraction depends on the wavelength of the light. The table in Equation 2 is for yellow light with a wavelength of 589 nm. The index depends on the wavelength because the speed of light in a material depends on its wavelength. The speed of light in any material is slower for shorter wavelengths, so the index of refraction is greater for light of shorter wavelengths. For example, for a type of glass called crown glass, $n = 1.50$ for red light, and $n = 1.53$ for violet light, which has a shorter wavelength than red light.

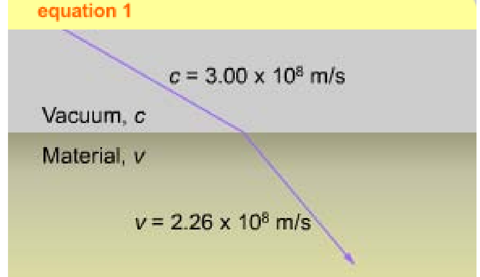
The index of refraction of a material also depends somewhat on its temperature. For certain crystalline materials, it also depends on the angle at which the light travels through the crystal lattice. This is why the index of refraction for corundum is indicated as approximate in the table.

The index of refraction for visible light in the Earth's atmosphere is about 1.0003 at standard temperature and pressure. Since we typically use two or three significant figures, we treat that index as 1.00. At the same level of precision, the speed of light is 3.00×10^8 m/s.

concept 1

Index of refraction
Speed in vacuum / speed in material

equation 1



Index of refraction

$$n = \frac{c}{v}$$

n = index of refraction
c = speed of light in vacuum
v = speed of light in material

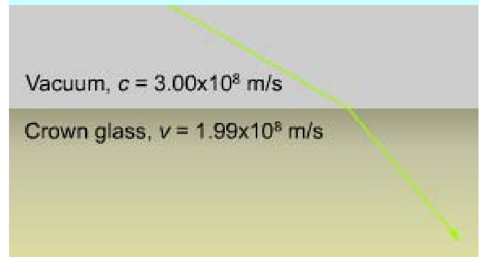
equation 2

	Index of refraction
Air	1.0003
Water	1.33
Vegetable oil	1.47
Crown glass	1.51
Salt	1.54
Flint glass	1.61
Corundum (ruby, sapphire)	1.77*
Diamond	2.42

At 20° C, λ = 589 nm *Approximate value

Indices of refraction

example 1



Green light travels at 1.99×10^8 m/s in crown glass. What is the index of refraction of the glass for this light?

$$n = \frac{c}{v}$$

$$n = \frac{3.00 \times 10^8 \text{ m/s}}{1.99 \times 10^8 \text{ m/s}}$$

$$n = 1.51$$

36.3 - Snell's law

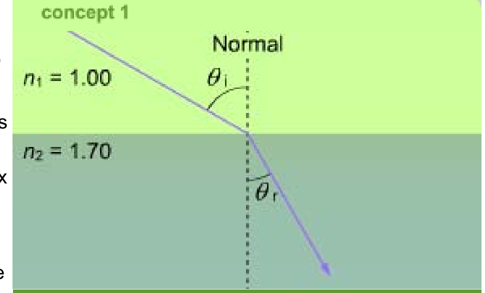
Snell's law is used to quantify refraction. In doing so, it uses some of the same terminology as the law of reflection. As with reflection, angles are measured between a ray and a line normal (perpendicular) to a surface. You see this illustrated in Concept 1, with both the angle of incidence (θ_i) and the angle of refraction (θ_r) shown.

Snell's law, shown in Equation 1, expresses the relationship between these angles. This law was discovered empirically by Willebrord Snell and written in its current form by René Descartes. It states that the product of the sine of the incident angle and the index of refraction of the incident medium equals the product of the sine of the refraction angle and the index of refraction of the refracting medium.

We also show the same equation in an alternate formulation: The ratio of the sine of the incident angle to the sine of the refracted angle is the **reciprocal** of the ratio of the first to the second index of refraction.

To put it more concretely, light bends **toward** the normal when it slows down, for instance, when it passes from air to water. You see this in Concept 1 to the right. It bends **away** from the normal when it speeds up, as from water to air.

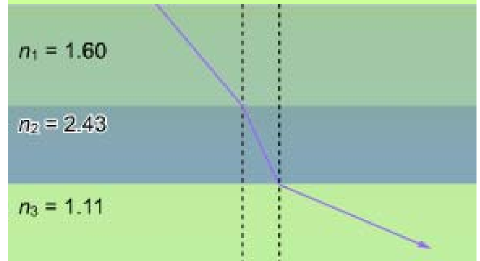
concept 1



Snell's law
Quantifies refraction
Slower light bends toward normal

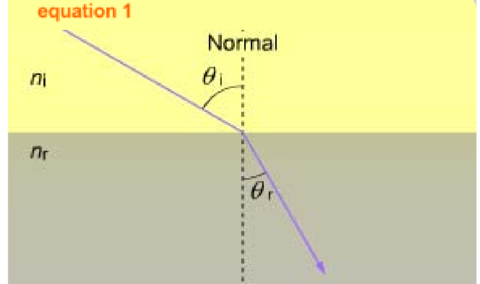
If light passes through several media, Snell's law can be applied at each interface. You see this occurring in Concept 2. The light's direction changes toward the normal at the first interface as it slows, and then away from the normal as it crosses the second interface. It bends away because light moves faster in the third medium, which has a lesser index of refraction than the second.

concept 2



Multiple media
Snell's law is applied multiple times

equation 1



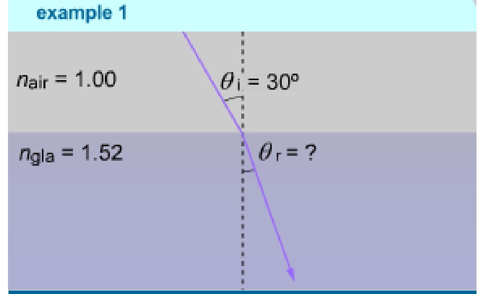
Snell's law

$$n_i \sin\theta_i = n_r \sin\theta_r$$

n = index of refraction
θ_i = incident angle from normal
θ_r = refracted angle from normal

$$\frac{\sin\theta_i}{\sin\theta_r} = \frac{n_r}{n_i}$$

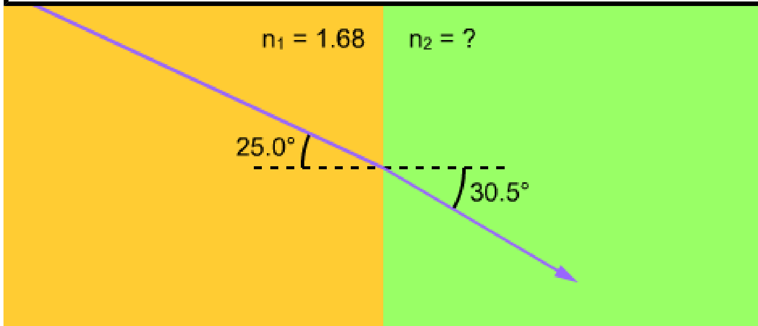
example 1



What is the angle of refraction here?

$n_{air} \sin \theta_i = n_{gla} \sin \theta_r$
$\sin \theta_r = (n_{air} \sin \theta_i) / n_{gla}$
$\theta_r = \arcsin [(n_{air} \sin \theta_i) / n_{gla}]$
$\theta_r = \arcsin [(1.00 \sin 30°)/1.52]$
$\theta_r = 19°$

36.4 - Interactive checkpoint: Snell's law



What is the index of refraction of the second material?

Answer:

n_2 =

36.5 - Everyday effects of refraction

The refraction of light can cause interesting and sometimes confusing results as the brain interprets the position of objects it sees via refracted light. For instance, in the upper illustration to the right, you see an ancient Egyptian fisherman trying to spear a fish. The solid line indicates the refracted path of the light traveling from the fish to his eyes. Since his brain expects light to travel in a straight line, he projects the fish to be in the position indicated at the end of the dashed line, which causes him to think the fish is nearer to the surface than it is. Experienced spear fishermen know how to compensate for this effect.

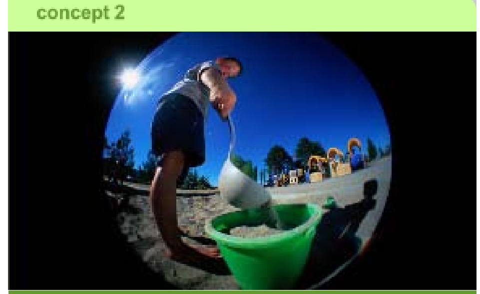
The fisherman is not the only one to experience the effects of refraction. The fish does, too. The light from the world above seen by the fish also refracts. Light coming straight down will pass through the water's surface unchanged, but light at any angle will refract, in the process giving the fish a wider field of view of the world above than it would have if there were no refraction. In essence, the fish sees a compressed wide-angle view of the scene above. Certain camera lenses, appropriately called fisheye lenses, can create the same effect, as illustrated in Concept 2.

Another interesting consequence of Snell's law concerns light passing through a window. The light refracts as it travels from air to glass. After passing through the glass, it refracts again at the second interface. The ratio of the indices of refraction is now reversed, so the initial change in angle is cancelled out. The light ray that emerges is parallel to the initial ray, but displaced a small amount. The amount of displacement is small enough that we ordinarily do not notice "window shift." However, if you place a newspaper page on a tabletop and cover half of it with a flat pane of glass, you will be able to observe the displacement effect.

concept 1

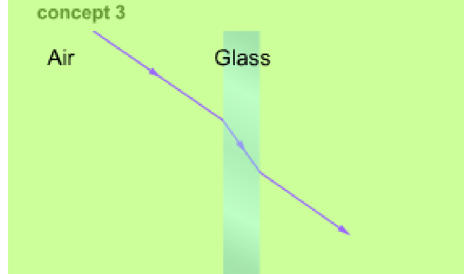
Refraction confuses fisherman
He misjudges position of fish

concept 2

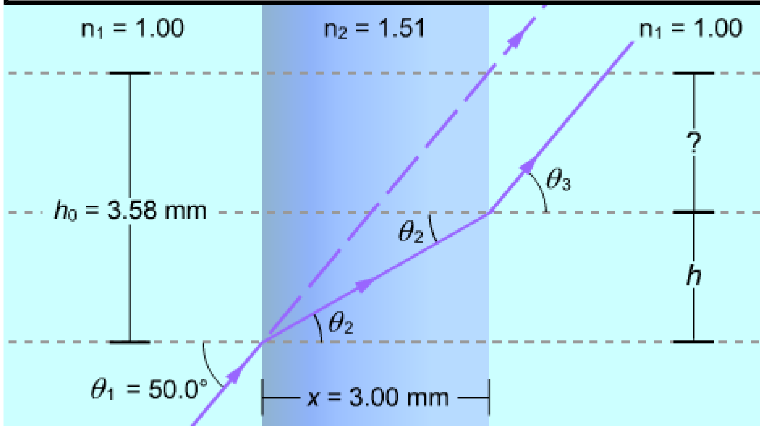


A fisheye view
Wide-angle view compresses images

concept 3



Light through a windowpane
Light refracts twice
Emerges parallel to initial ray



Show that a light ray exits a pane of glass parallel to the incident ray.

Find the vertical difference, on the right side of the glass, between the original unrefracted path of the ray and the actual doubly refracted ray.

There is no net change in direction of light passing through a pane of ordinary window glass, since its two surfaces are parallel. However, a light ray that strikes the glass at an oblique angle does refract twice, and it leaves the glass at a different location than it would if there were no refraction.

In the diagram above, this shift equals the vertical difference $h_0 - h$, where h_0 is the vertical distance an unrefracted ray would travel over a horizontal distance of 3.00 mm, and h is the vertical distance traveled by the actual refracted ray in the glass. With trigonometry it can be shown that h_0 is 3.58 mm. Your job is to show that the doubly refracted ray travels in the same direction as the incident ray, and to find the difference $h_0 - h$.

Variables

index of refraction of air	$n_1 = 1.00$
index of refraction of glass	$n_2 = 1.51$
incident angle	$\theta_1 = 50.0°$
intermediate angle	θ_2
exit angle	θ_3
vertical displacement without refraction	$h_0 = 3.58$ mm
vertical displacement with refraction	h
thickness of glass	$x = 3.00$ mm

Strategy

1. Show that the incident and exiting rays are parallel, using Snell's law.
2. Use Snell's law to calculate θ_2.
3. Use trigonometric relationships to solve for h.
4. Finally, calculate the shift in the refracted light as $h_0 - h$.

Physics principles and equations

Snell's law

$$n_i \sin \theta_i = n_r \sin \theta_r$$

Step-by-step solution

First we show that the incident and final rays are parallel.

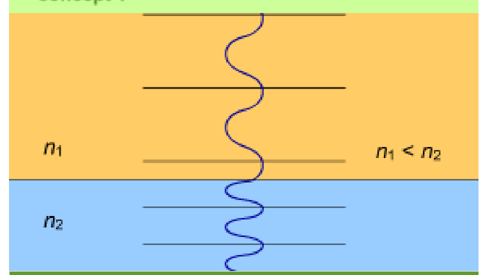
Step	Reason
1. $n_1 \sin \theta_1 = n_2 \sin \theta_2$	Snell's law
2. $n_2 \sin \theta_2 = n_1 \sin \theta_3$	Snell's law
3. $n_1 \sin \theta_1 = n_1 \sin \theta_3$	substitute equation 2 into equation 1
4. $\sin \theta_1 = \sin \theta_3$	simplify
5. $\theta_1 = \theta_3$	take arcsin of both sides

The equal incident and final angles show that the incident and exiting rays are parallel. Now we calculate the vertical displacement h of the doubly-refracted ray, and finally the difference $h_0 - h$.

Step	Reason
6. $\theta_2 = \arcsin(n_1 \sin\theta_1 / n_2)$	solve equation 1 for θ_2
7. $\theta_2 = \arcsin(1.00 \sin 50.0° / 1.51)$ $\theta_2 = 30.5°$	evaluate
8. $\tan\theta_2 = h / x$	trigonometry
9. $h = x \tan\theta_2$	solve for h
10. $h = (3.00\text{ mm}) \tan 30.5°$ $h = 1.77$ mm	evaluate
11. $h_0 - h = 3.58\text{ mm} - 1.77\text{ mm}$ $h_0 - h = 1.81$ mm	calculate difference

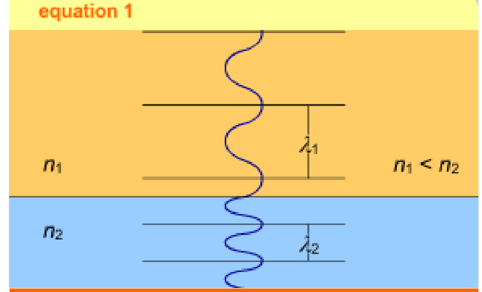
36.7 - Wavelength of light in different media

concept 1



Frequency and wavelength
When light changes speed
· The frequency stays the same
· The wavelength changes

equation 1



Wavelength, index ratios inversely proportional

$$\frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1}$$

λ = wavelength in medium
n = index of refraction

When light changes speed as it moves from one medium to another, its frequency stays the same but its wavelength changes. The ratio of its wavelengths in the two media is the inverse of the ratio of the indices of refraction. We show this as an equation to the right and derive it below.

Before deriving the equation, let's consider why the frequency stays the same, since this is an essential part of the derivation. The frequencies in the media must be the same, because if they were not, waves would either pile up at the interface or be destroyed. Neither occurs. You can witness this at the beach, where wave speed and wavelength may change as waves approach the beach, but the frequency of the waves does not change.

Variables

In this derivation, c represents the speed of light in a vacuum. For the other two media we define the variables in the following table:

	medium 1	medium 2
light speed	v_1	v_2
frequency	f_1	f_2
wavelength	λ_1	λ_2
index of refraction	n_1	n_2

Strategy

1. Use the equality of frequencies in the two media together with the wave speed equation to obtain a proportionality of the light speeds and wavelengths in the media.
2. Use the definition of the index of refraction to convert the previous proportion to one involving wavelengths and indices of refraction.

Physics principles and equations

The wave speed equation states that for any wave, the speed is the product of the wavelength and the frequency:

$$v = \lambda f$$

As a wave passes from one medium to another, its speed and wavelength may change, but its frequency must remain the same.

The definition of the index of refraction of a medium is

$$n = \frac{c}{v}$$

Step-by-step derivation

First we derive a proportion containing the light speeds and wavelengths of light in the two media.

Step	Reason
1. $f = \dfrac{v}{\lambda}$	wave speed equation
2. $f_1 = f_2$	constant frequency
3. $\dfrac{v_1}{\lambda_1} = \dfrac{v_2}{\lambda_2}$	substitute 1 into 2

Now we convert the previous proportion into one containing the wavelengths and indices of refraction of light in the two media.

Step	Reason
4. $v = c/n$	definition of index of refraction
5. $\dfrac{c/n_1}{\lambda_1} = \dfrac{c/n_2}{\lambda_2}$	substitute 4 into 3
6. $\dfrac{\lambda_1}{\lambda_2} = \dfrac{n_2}{n_1}$	rearrange

Note that the ratio of wavelengths is inversely proportional to the ratio of indices of refraction. This means that, in a material with a larger index, the wavelength of light will be less.

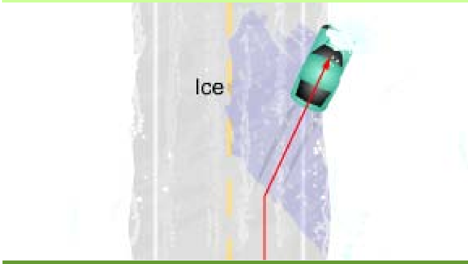
36.8 - Why refraction occurs

Why do waves change direction when they change speed? Here we offer a mechanical analogy to explain this phenomenon.

Consider the diagram in Concept 1 to the right. It uses the example of a car encountering ice on the side of a road. If this has ever happened to you, you know that the car can dangerously change direction when one (or more) of its wheels encounters ice, as shown in the diagram. Why is this so? The wheels on the ice provide less traction, so the right side of the car moves more slowly. The left side of the car continues to move at the same speed as before, which causes the car to rotate and veer off the road. When all the wheels supplying power are on the ice, the car will once again move straight ahead because both sides will be moving at the same speed.

Why light refracts can also be explained using Fermat's principle of least time. This principle, developed by French mathematician Pierre de Fermat (1601−1665), states that light will travel the path between two points that requires the least amount of time. This may seem like it should be a straight line, but it is not when the speed varies along the path between two points.

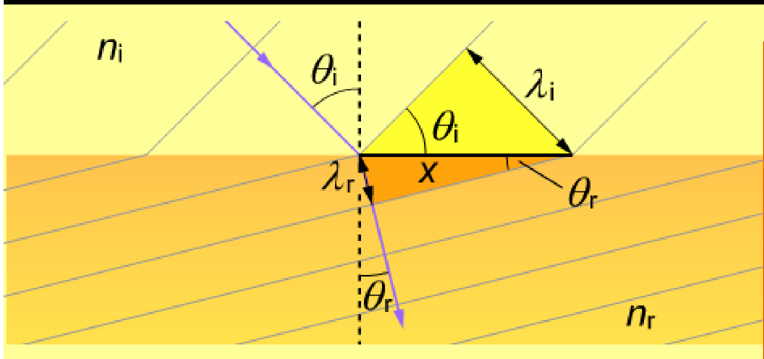
concept 1



Car hits ice
Speed change on one side changes direction of car

This timesaving "technique" is similar to what you would intuitively do if you were standing on a beach and saw a swimmer floundering desperately in the water some distance down the shore. Instead of taking a direct straight-line path to the rescue, you would run along the beach for some distance, counting on your greater land speed, before taking to the water.

36.9 - Derivation: Snell's law



equation 1

Snell's law

$$\frac{\sin\theta_i}{\sin\theta_r} = \frac{n_r}{n_i}$$

θ_i = incident angle
θ_r = refracted angle
n = index of refraction

We can derive Snell's law by analyzing wave fronts as they move across media with differing indices of refraction. To derive the law, we use the fact that the wavelength of light changes as it moves across these media. The derivation relies on the geometry associated with this change in wavelength.

First, we explain the diagram you see above. The purple line is a light ray refracting at an interface. In the diagram, light travels more slowly in the lower medium than the upper. This could represent, for example, light passing from air into water. The gray lines perpendicular to the ray represent wave fronts. You see the wavelength labeled as λ (λ_i in the upper medium, λ_r in the lower medium).

There are two right triangles in the diagram that share the hypotenuse labeled x. The bright yellow triangle shows elements of a wave front that has not yet entered the lower medium. The dark orange triangle shows elements of a wave front that is now traveling in the lower medium. The angles of incidence and refraction θ_i and θ_r are also shown in the diagram. Because the wave fronts are perpendicular to the light rays, we can identify angles in each of the triangles that are equal to θ_i and θ_r. These base angles are shown in the diagram.

Variables

In this derivation, x represents the common hypotenuse of the two triangles in the diagram. For the incident and refractive media we define the variables in the following table.

	incident medium	refractive medium
angle	θ_i	θ_r
wavelength	λ_i	λ_r
index of refraction	n_i	n_r

Strategy

1. Consider the two triangles in the diagram. State the sines of their base angles θ_i and θ_r as trigonometric ratios of the triangles' sides.
2. Construct the ratio $\sin\theta_i / \sin\theta_r$. The common hypotenuse x will cancel out, leaving a ratio of wavelengths.
3. Restate the ratio of wavelengths as a ratio of indices of refraction to obtain Snell's law.

Physics principles and equations

The ratio of the wavelengths is inversely proportional to the ratio of the indices of refraction.

$$\frac{\lambda_i}{\lambda_r} = \frac{n_r}{n_i}$$

Step-by-step derivation

We construct the fraction $\sin\theta_i / \sin\theta_r$, and calculate the sines as the ratios of the sides of triangles. This leads to a ratio of wavelengths that can be replaced by a ratio of indices of refraction, yielding Snell's law.

Step	Reason
1. $\sin\theta_i = \lambda_i / x$, $\sin\theta_r = \lambda_r / x$	definition of sine
2. $\dfrac{\sin\theta_i}{\sin\theta_r} = \dfrac{\lambda_i/x}{\lambda_r/x}$	ratio using definition of sine
3. $\dfrac{\sin\theta_i}{\sin\theta_r} = \dfrac{\lambda_i}{\lambda_r}$	simplify
4. $\dfrac{\lambda_i}{\lambda_r} = \dfrac{n_r}{n_i}$	change of wavelength
5. $\dfrac{\sin\theta_i}{\sin\theta_r} = \dfrac{n_r}{n_i}$	substitute equation 4 into equation 3

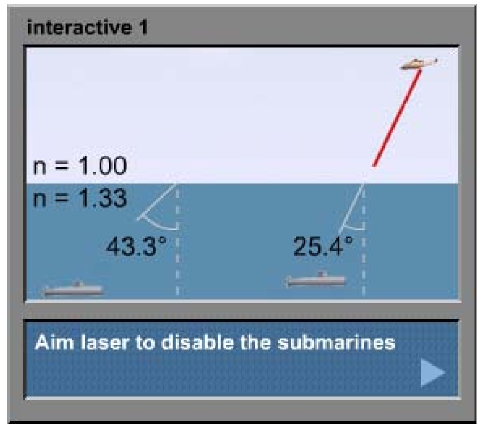
36.10 - Interactive problem: helicopter and submarines

The simulation on the right is similar to the one in this chapter's introduction. As before, submarines lurk under the waves, but now you have only a single helicopter.

As before, you have a laser you can aim in an attempt to disable two submarines before they disable you. You keep shooting as long as you keep making hits. To shoot your laser, aim it by dragging the aiming arrow and fire it by clicking on FIRE. The angle of incidence is shown in an output gauge.

As soon as you miss, it is the submarines' turn. Warning: The computer has been set to be far more accurate in this game than in the introductory one. Unless you are very precise with your shots, it is unlikely you will win.

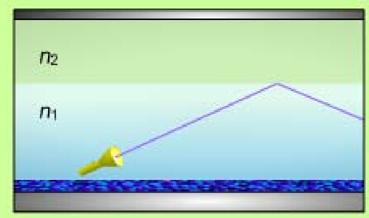
However, there is good news: Now you have more intellectual firepower because you have the aid of Snell's law. You are also given some assistance from an able comrade; she has computed the angles of refraction required for your laser to reach the submarines, as shown in the diagram. You should use 1.33 for the index of refraction of water, and 1.00 for the index of refraction of air. If you correctly set the angle of incidence when you aim each of your shots, you can make two straight hits and disable the submarines before they disable you.



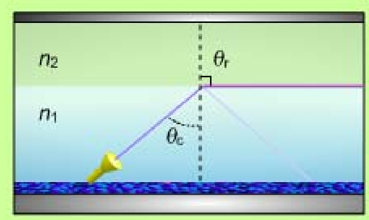
36.11 - Total internal reflection

Total internal reflection: Light reflects completely at an interface, back into the medium with the higher refractive index.

Critical angle: The minimum angle of incidence at which total internal reflection occurs.



Total internal reflection
No incident light leaves initial medium
Light is reflected at interface



Critical angle
Minimum incident angle for total internal reflection

Total internal reflection means no light passes from one medium to another. On the right, we show how this occurs using the example of an underwater flashlight shining a beam at the interface between water and air. As the diagram shows, all the light reflects back into the medium with the higher refractive index, the water. No light passes into the air above the water.

Why does this occur? Consider what happens when light is directed from water into air. Water has a greater index of refraction than air. At relatively small angles of incidence, light passes from the water into the air, and as it does so, it refracts away from the normal. The angle of refraction is greater than the angle of incidence.

As the angle of incidence increases, the angle of refraction will increase as well, and it will always be greater than the angle of incidence. At a sufficient angle of incidence, the angle of refraction reaches $90°$, perpendicular to the normal and parallel to the surface of the water. The light no longer crosses the interface but remains in the water. The minimum angle of incidence at which this occurs is called the critical angle.

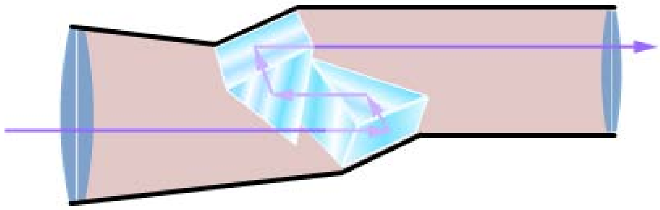
Even at angles of incidence less than the critical angle there is some internal reflection along with refraction. But when the critical angle is exceeded, there is no refraction at all: All of the light is internally reflected.

Total internal reflection is shown in Concept 1. All the light is reflected back into the water. Concept 2 shows the situation when the angle of incidence is equal to the critical angle. The critical angle depends on the ratio of the indices of refraction of the two media. You see this as the equation on the right, which is obtained by setting $\theta_r = 90°$ (so $\sin\theta_r = 1$) in Snell's law.

Diamonds rely on total internal reflection to achieve their sparkle. They are cut so that light entering them reflects internally and emerges only at certain points, giving the effect of scintillating light.

If you happen to have a diamond ring handy, you can demonstrate this dependence on the ratio of indices by placing the ring in water. The sparkle disappears; this is because the indices of refraction of diamond and water do not differ as significantly as those of diamond and air, so the geometry of the ring no longer causes the same internal reflection.

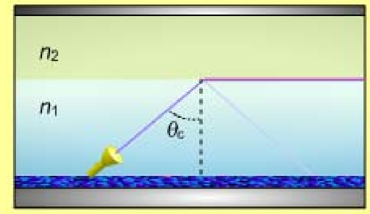
Engineers use total internal reflection in many different applications. For instance, right isosceles prisms are used in binoculars to redirect light, as in the illustration below.



Fiber optic cable is used to transfer information (in glass or plastic) using total internal reflection. Data (such as speech) is first encoded as modulations of a beam of laser light. The light remains inside the transparent cable as it travels due to its total internal reflection off the inner surfaces of the cable walls. In this way, light is transmitted through the cable with little loss.

Using light instead of electricity to transfer information has several benefits. Light can be used to encode much more information than electric oscillations because of its extremely high frequency (more than 10^{14} Hz for red light), and fiber optic cable is immune to interference problems from nearby electrical applications.

equation 1



Critical angle

$$\sin\theta_c = \frac{n_2}{n_1}$$

θ_c = critical angle
n_1, n_2 = indices of refraction ($n_1 > n_2$)

example 1

You are cutting a sapphire to make it as brilliant as possible. Find the critical angle for the sapphire in air. The index of refraction of a sapphire is 1.77.

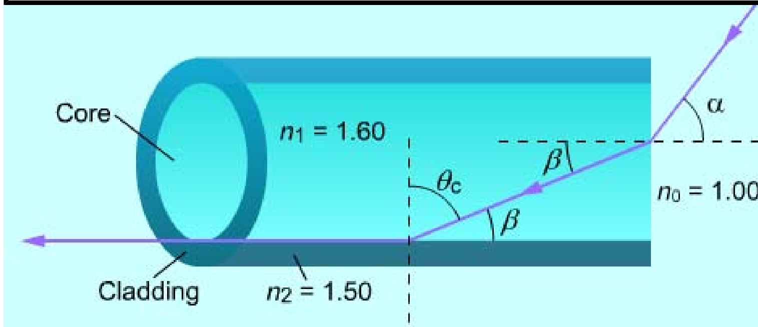
$$\sin\theta_c = \frac{n_{air}}{n_{sap}}$$

$$\theta_c = \arcsin\left(\frac{n_{air}}{n_{sap}}\right)$$

$$\theta_c = \arcsin\left(\frac{1.00}{1.77}\right)$$

$$\theta_c = 34.4°$$

36.12 - Sample problem: fiber optic cable



What is the maximum angle α that results in total internal reflection inside the fiber optic cable?

Above is a diagram of a fiber optic cable. A fiber optic cable consists of two transparent substances having different indices of refraction. The core has a higher index of refraction than the cladding that surrounds it. Because of this, light rays can zigzag down the core for a very long distance, reflecting off the inside surface of the cladding by total internal reflection within the core.

Light is fed into the cable at one end by a signal-generating device. Fiber optic engineers use an angle called the *acceptance angle*, labeled above in the diagram as α. This is the maximum angle at which a light ray can enter a fiber optic cable and be transmitted along it by total internal reflection. As the light enters the cable at a greater angle α, the angle β becomes greater as well, which means the angle of incidence at the core-cladding interface becomes smaller. If the angle of incidence is less than the critical angle, light "leaks" out of the cable.

As an aside: Fiber optic cables are flexible. If they are bent too sharply, light leaks out. Look at the cable above, and imagine the left side being raised so the cable bends. This reduces the angle of incidence at the core-cladding interface, and if it is reduced too much, it becomes less than the critical angle.

Variables

index of refraction of air	$n_0 = 1.00$
index of refraction of core	$n_1 = 1.60$
index of refraction of cladding	$n_2 = 1.50$
acceptance angle	α
intermediate angle	β
critical angle	θ_c

Strategy

1. The angles β and θ_c in the diagram are complementary. Use trigonometric identities to establish a relationship between $\sin\beta$ and $\sin\theta_c$.
2. Use Snell's law to write $\sin\alpha$ in terms of $\sin\beta$. By employing the result of previous steps, write $\sin\alpha$ in terms of $\sin\theta_c$.
3. Finally, use the critical angle formula to find $\sin\alpha$ in terms of n_0, n_1, and n_2. Use the given values in the problem to evaluate α.

Physics principles and equations

Snell's law

$$n_i \sin\theta_i = n_r \sin\theta_r$$

Critical angle formula

$$\sin\theta_c = n_2 / n_1$$

Mathematics principles

For complementary angles θ and φ,

$$\sin\theta = \cos\varphi.$$

Trigonometric identity

$$\cos^2\theta + \sin^2\theta = 1$$

The two angles labeled β in the diagram are equal because the normal line at the end of the cable is parallel to the sides of the cable.

Step-by-step solution

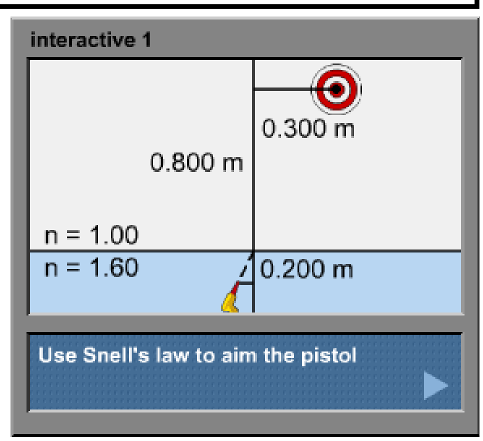
We start by using trigonometry to establish a relationship between $\sin\beta$ and $\sin\theta_c$.

Step	Reason
1. $\cos\beta = \sin\theta_c$	trigonometric identity
2. $\sin\beta = \sqrt{1-\cos^2\beta}$	trigonometric identity
3. $\sin\beta = \sqrt{1-\sin^2\theta_c}$	substitute equation 1 into equation 2

In the following steps we use Snell's law and the definition of the critical angle to find $\sin\alpha$ in terms of indices of refraction. Then we evaluate to find a value for the acceptance angle α.

Step	Reason
4. $n_0 \sin\alpha = n_1 \sin\beta$	Snell's law
5. $\sin\alpha = \frac{n_1}{n_0}\sqrt{1-\sin^2\theta_c}$	substitute equation 3 into equation 4
6. $\sin\theta_c = \frac{n_2}{n_1}$	definition of critical angle
7. $\sin\alpha = \frac{n_1}{n_0}\sqrt{1-\frac{n_2^2}{n_1^2}}$	substitute equation 6 into equation 5
8. $\sin\alpha = \frac{1.60}{1.00}\sqrt{1-\frac{(1.50)^2}{(1.60)^2}}$ $\sin\alpha = 0.557$ $\alpha = 33.8°$	evaluate

36.13 - Interactive problem: laser target pistol



In this simulation, you fire a laser pistol at two stationary targets. You want a bulls-eye, a shot that passes through the central black circle of the target.

You happen to be immersed in a fluid with an index of refraction of 1.60. When you shoot the laser from within the fluid, you have to account for the refraction of your laser beam as it passes from the fluid into the air. Some portion of the beam also reflects from the interface. If the incident angle is equal to or greater than the critical angle, all of the beam will reflect.

The laser pistol is designed so its beam always hits the interface between the fluid and the air at the same point. You drag the pistol from side to side to determine its angle of incidence. An output gauge tells you the horizontal distance between the pistol and where its beam passes through the interface. You are also told how far the pistol is below the surface, a value that stays constant.

There are two targets to hit: one above the surface of the fluid and the other below the surface. You will have to use a different strategy to hit each target. For each target you know its horizontal distance from where the laser passes through the interface, and its height above the fluid. You must strike each target in its center to succeed.

You have just read a fair number of facts. It may be useful now to look at the diagram and the simulation. You will probably find the lower target easier to hit, if you just reflect a bit. You can then bend your mind toward determining how to hit the upper target. A good way to start is to apply trigonometry to calculate the necessary angle of refraction. Then, using Snell's law and some more trigonometry, you can determine where to position the pistol. When you have calculated the correct horizontal distance, drag the pistol and press FIRE to test your answer.

A rainbow provides an example of dispersion.

Dispersion: Refraction that causes light to separate into its various wavelengths.

Humans perceive different wavelengths of light as different colors. When we see light of a single wavelength, we perceive it as a pure color such as red, green or violet. The light that comes from the Sun and from standard light bulbs, which we perceive as white light, consists of a mixture of many different wavelengths of electromagnetic radiation.

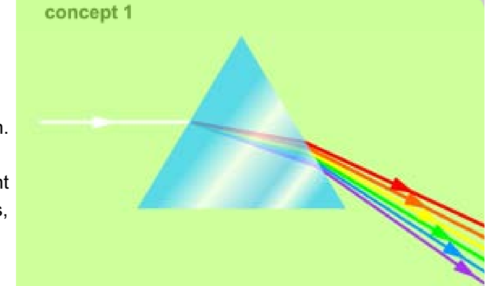
Prisms, such as the one shown to the right, separate white light into its many wavelengths. This is called dispersion, or to be more specific, *chromatic dispersion*, since dispersion can also occur with non-visible electromagnetic radiation and mechanical waves as well. Sir Isaac Newton famously used a pair of prisms to disperse white sunlight into colors and recombine the colors into white light.

Prisms disperse light because the index of refraction depends on the light's wavelength. The index of refraction for a given material is greater for waves of shorter wavelength, so blue and violet light refract more than red or orange light. When white light is incident upon a prism, the different wavelengths that make it up are refracted at different angles, resulting in a rainbow of colors.

The rainbows we see in the sky are also caused by dispersion. Light disperses as it enters a spherical raindrop. Total internal reflection occurs inside the raindrop, and the light then refracts again when leaving the drop, dispersing even more. You see this illustrated in a raindrop to the right.

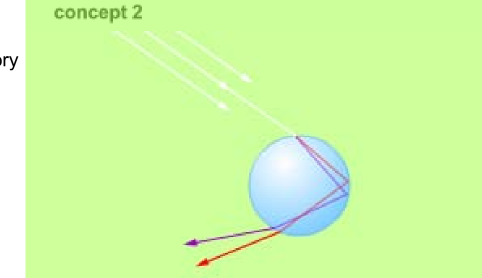
Why does this process cause rainbows? Raindrops reflect back nested "cones" of light of different colors. In Concept 3 you see how red light reflected back from every drop in the outermost portions of the rainbow reaches the observer, but the violet light from those drops passes him by. You also see how the violet light from the innermost portions of the rainbow reaches him, but not the red light. Other topics in rainbow theory are still being researched. For instance, physicists, although vaguely alluding to quantum effects and the promise of string theory, have to date offered no convincing explanation for the pot of gold at the end of the rainbow.

concept 1



Dispersion
Creation of a spectrum by refraction
Different wavelengths refract differently

concept 2



Rainbow: inside a raindrop
Refraction plus total internal reflection
Colors refract at different angles

concept 3

Observer views rainbow
Red light from outer drops reaches him
Violet light from inner drops reaches him

36.15 - Gotchas

You calculated a material as having a refractive index of 0.76. Oops! This is not possible. The speed of light in a vacuum is greater than its speed in any medium, meaning that the refractive index of the material must be greater than 1.00. Remember that when calculating the index of refraction, you put the speed of light in a vacuum in the numerator, not the denominator.

Light strikes water in a placid lake at an angle of $30°$ from the horizontal, so its incident angle is $30°$. No, the incident angle is $60°$. It is measured between the incident ray and the interface's **normal line**. This line is perpendicular to the interface.

Light passes from one medium to another, and bends toward the normal. The refractive index of the second material is greater than the first. Yes, it is. You can confirm this using Snell's law.

All wavelengths (colors) of light refract at the same angle when crossing the interface between different media. No, the angle of refraction varies by wavelength.

36.16 - Summary

Refraction is the changing of a wave's direction due to a change in its speed. It occurs when a wave, like light, passes from one medium into another.

A material's index of refraction is the ratio of the speed of light in a vacuum (c) to its speed in the material. The index of refraction is represented by the letter n.

Snell's law quantifies how much light bends when it crosses an interface. Light bends towards a line normal to the interface between the media as it slows down (passes into a medium with a higher index of refraction). When light passes through multiple media its path can be calculated by applying Snell's law at each interface.

Because light waves do not accumulate or disappear at an interface between two media, the frequency of light stays the same as it passes from one material to another. Because the speed of the light changes, the wavelength must change.

Total internal reflection occurs when no light is refracted at an interface; it is all reflected. This is only possible as light encounters an interface with a medium having a smaller index of refraction, at a large enough incident angle. The critical angle is the minimum angle of incidence at which total internal reflection occurs.

Equations

Index of refraction

$n = c/v$

Snell's law

$n_i \sin \theta_i = n_r \sin \theta_r$

Critical angle

$\sin\theta_c = \frac{n_2}{n_1}$

The index of refraction for a material varies somewhat by the wavelength of light being refracted. Light composed of many wavelengths, such as white light, can be separated into a spectrum by refraction. This effect is called dispersion. Prisms exhibit dispersion.

Chapter 36 Problems

Chapter Assumptions

Unless stated otherwise, the incident medium is air, with an index of refraction of 1.00.

Conceptual Problems

C.1 Light of wavelength X travels more slowly in glass than light of wavelength Y. (a) Which light refracts (bends) more as it travels from air to glass? (b) Which wavelength is longer?

(a) ○ X ○ Y
(b) ○ X ○ Y

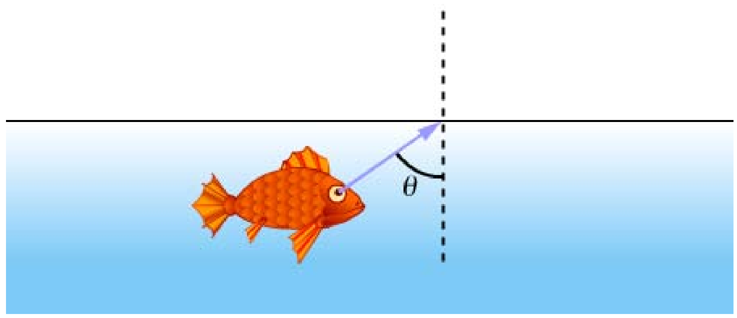
C.2 As light travels between media with different indices of refraction, which of the following wave properties changes: speed, frequency, wavelength?

☐ Speed
☐ Frequency
☐ Wavelength

C.3 Sparkle refers to percentage of light out to light in. Why do cut diamonds, whose index of refraction is 2.42, sparkle more than identically cut glass with an index of refraction of 1.60? Diamonds are cut so that a maximum of the light entering the diamond's top eventually leaves through the top.

C.4 If a fish in a still river tries to look up out of the water at an angle (measured from the vertical) that is greater than the critical angle, will he be able to see the world of air? If not, what will he see?

○ Yes ○ No



Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. Consider the smaller angles between the laser's path and the dashed normal line. (a) Which is larger, the angle in the air, or the angle in the water? (b) When you shoot a laser from a helicopter toward a submarine, does the laser bend toward or away from the dashed normal line as it changes media?

(a) ○ Air ○ Water
(b) ○ Toward ○ Away from

Section 2 - Index of refraction

2.1 A laser pulse travels directly through a flint glass plate that is 0.150 mm thick, and then through a diamond sliver that is 0.130 mm thick. In which substance does the pulse spend more time?

○ Flint glass
○ Diamond

2.2 Yellow light travels 0.680 times as fast in a certain type of plastic than in a vacuum. What is the index of refraction for yellow light in the plastic?

____

2.3 At what speed does light of wavelength 589 nm travel in flint glass? Flint glass has an index of refraction of 1.61 for that wavelength.

__________ m/s

2.4 Red light travels at 2.07×10^{8} m/s in a particular type of oil. What is the index of refraction for red light in the oil?

____

Section 3 - Snell's law

3.1 A beam of light travels from a medium with an index of refraction of 1.25 to a medium with an index of refraction of 1.46. If the incoming beam makes an angle of 14.0° with the normal, at what angle from the normal will it refract?

_________ °

3.2 A beam of light travels from a medium with an index of refraction of 1.65 to a medium with an index of refraction of 1.30. If the beam refracts at an angle of 23.5° from the normal, at what angle to the normal did it enter the second medium?

_________ °

3.3 A beam of light crosses the interface between two media with an angle of incidence of 34.0° and an angle of refraction of 24.6°. The second medium has an index of refraction of 1.70. What is the index of refraction of the first medium?

____

3.4 A double-paned window consists of two parallel panes of glass with an air gap between them. The outside pane has an index of refraction of 1.61 while the inside layer has an index of refraction of 1.51. If light enters the window from outside at 21.0° to the normal, at what angle to the normal does it exit the window on the inside?

_________ °

3.5 A beam of light travels from one medium to another, with the incident beam making an angle of 19.0° from the normal. The beam then refracts at an angle of 32.0°. How many times faster does the light travel in the second medium than the first?

____

3.6 A glass of water (n = 1.33) has a layer of vegetable oil (n = 1.47) floating on top. A beam of light enters the oil from above the glass with an incident angle of 57.0°. What is the refraction angle of the light leaving the oil-water interface?

_________ °

3.7 A ring consists of a diamond (n = 2.42) and another jewel mounted side by side on a flat plate. Their top surfaces are also parallel to the plate. When laser light is aimed at the ring at 80.0° to the normal, the refraction angle in the diamond is 11.1° less (from the normal) than the refraction angle in the other jewel. What is the index of refraction of this other jewel?

____

Section 7 - Wavelength of light in different media

7.1 A beam of light crosses an interface between two media, its wavelength increasing by a factor of 1.40. If the second medium has an index of refraction of 1.37, what is the index of refraction of the first medium?

____

Section 10 - Interactive problem: helicopter and submarines

10.1 Use the simulation in the interactive problem in this section to calculate the angles of incidence required to hit the (a) left and (b) right submarine.

(a) _________ °
(b) _________ °

Section 11 - Total internal reflection

11.1 Find the critical angle for a ray of yellow light traveling from diamond (n = 2.42) into air (n = 1.00).

_________ °

11.2 The critical angle for light traveling from a certain plastic into air is 35.5°. What is the index of refraction of the plastic?

____

11.3 You are constructing a fiber optic cable. You wish to surround the core of the cable with a cladding material such that the critical angle to go from the core to the cladding is 73.0°. If the index of refraction of the core material is 1.61, what must the index of refraction of the cladding material be?

____

11.4 A point source of light is at the bottom of a koi pond, at a depth of 0.525 meters. What is the radius of the circle of light formed on the water's surface? Take the index of refraction of water to be 1.33. Hint: Some of the light emitted experiences total internal reflection inside the water.

_________ m

Section 13 - Interactive problem: laser target pistol

13.1 Use the simulation in the interactive problem in this section to calculate the horizontal distance required for the laser pistol to hit the (a) top and (b) bottom targets.

(a) __________ m
(b) __________ m

Section 14 - Dispersion and prisms

14.1 A red ray and a blue ray each pass from air into silica at an incident angle of 29.00°. How large is the angle between the two refracted rays? Red light has an index of refraction in silica of 1.459 and blue light has an index of 1.467.

__________ °

14.2 A narrow ray of white light traveling in air enters a glass block at an incident angle of 40.00°. The white light consists of a range of wavelengths. For these colors, the indices of refraction for the glass range from 1.512 to 1.531. Find the angular width of the ray inside the block.

__________ °

14.3 You have a glass prism shaped like an equilateral triangular slab. A yellow light ray, for which the index of refraction is 1.50, enters one face at an incident angle of 73.0° to the normal. (a) What is the angle of refraction when the light first enters the prism? (b) What is the angle of refraction when the light leaves the prism? (c) What is the angle between the ray that enters the prism and the one that leaves?

(a) __________ °
(b) __________ °
(c) __________ °

Additional Problems

A.1 The high-tide marker in a bay stands vertically with its lower 0.560 meters submerged. Its total height is 2.50 meters. If the Sun is currently 30.0° above the horizon, what is the length of the marker's shadow on the bottom of the bay, assuming the bottom is flat? Use 1.33 as the index of refraction of the water.

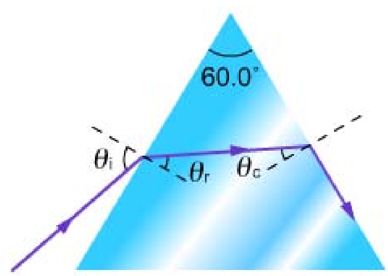
__________ m

A.2 A vertical timber, part of a pier that has decayed, is completely submerged in water. The timber is 1.50 meters tall. If the Sun is 30.0° above the horizon, what is the length of the timber's shadow? The index of refraction of water is 1.33.

__________ m

A.3 You have a glass prism shaped like an equilateral triangular slab as in the sketch. What is the minimum incident angle, θ_i, at which a light ray can enter the prism so that it does not undergo total internal reflection upon exiting the prism? Assume that the index of refraction is 1.51. State your answer to the nearest degree.

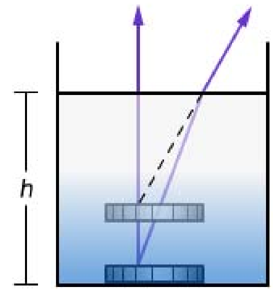
__________ °



A.4 A light ray enters the center of one end of a thin cylinder of glass at an incident angle of 55.0°. The cylinder has a diameter of 4.30 mm and a length of 595 mm. How many times will the light ray reflect off the inside surface of the cylinder before exiting? The glass has an index of refraction of 1.50.

__________ times

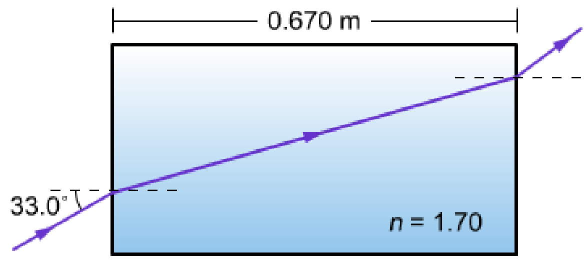
A.5 A coin is placed at the bottom of a glass, which is then filled with a transparent liquid. The illustration shows two light rays coming from the coin. When the coin is viewed from above the liquid looking down almost normal to the surface, the point where the light rays appear to intersect is the apparent depth of the coin. Air above the liquid has an index of refraction of 1.00. If the coin appears to be 2/3 as deep as it actually is, what is the index of refraction of the liquid? Hint: For small angles sin $\theta \approx$ tan θ.



____

A.6 A light ray enters the plastic block in the illustration at an angle of 33.0° and then exits the opposite side. The long side of the block has a length of 0.670 m and its index of refraction is 1.70. For how much time is the light ray inside the block?

__________ s

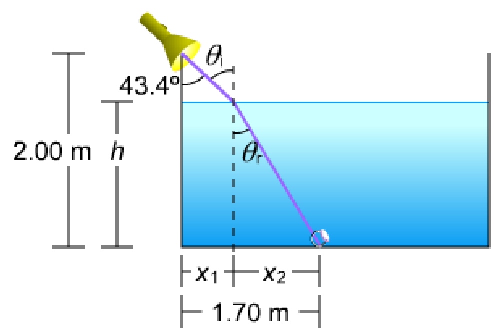


A.7 Sara is taller than Robin. Robin, whose eyes are 1.73 m above the ground, looks through a very thick window at Sara. A light ray coming from the top of Sara's head strikes the window at an angle of 25.0° from the normal, refracts when it enters the glass, refracts again when it exits the glass, and then enters Robin's eye. The window is 0.250 m thick and has an index of refraction of 1.61. Robin and Sara are each standing 0.100 m away from it. How tall is Sara?

__________ m

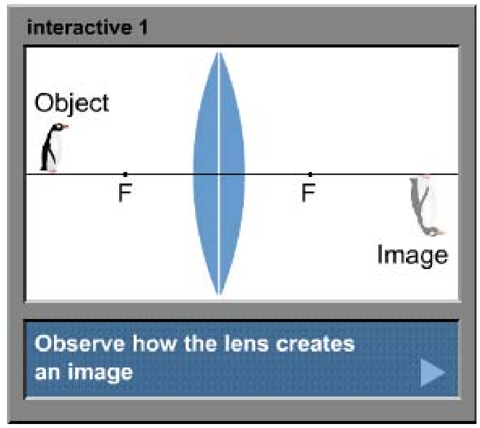
A.8 A flashlight beam shines into a tank of water (n = 1.33). The flashlight is positioned 2.00 m up the side of the tank, and the beam makes a 43.4° angle with the side of the tank. A ring is on the bottom of the tank, 1.70 m from the side. What is the height of the water level in the tank if the beam shines directly onto the ring? The drawing is not to scale.

__________ m



37 Lenses

37.0 - Introduction



Lenses may be the application of refraction with which you are most familiar. Lenses of glass or plastic, fashioned into eyeglasses and contact lenses, are worn by about 60% of the population of the United States, and this percentage is projected to increase as the population ages. Even people who do not require visual correction are "wearing" lenses anyway, since the eye itself contains an organic lens.

The human eye has the ability to refract light and focus it on the retina at the back of the eye, where signals are then transmitted to the brain. If the focusing is not precise enough, glasses or contact lenses can improve the results. These days, the eye itself can be tuned via laser surgery as well.

Lenses augment vision in other fashions as well. Lenses in magnifying glasses, telescopes, and microscopes allow you to see things that would otherwise be too small or too far away to see in detail.

At the right, you can begin your experimentation with lenses. The convex lens will produce an image of the penguin. You can move the penguin left and right to see how its distance from the lens changes the resulting image. Where must the object be relative to the focal point F to create an image on the same side of the lens as the object? Where must it be to create an image on the far side? Can you make the image appear right-side up (rather than inverted, as it is in the picture on the right)? A note: When the object is very near the focal point, the image will be too far away from the lens to be shown in the simulation's window.

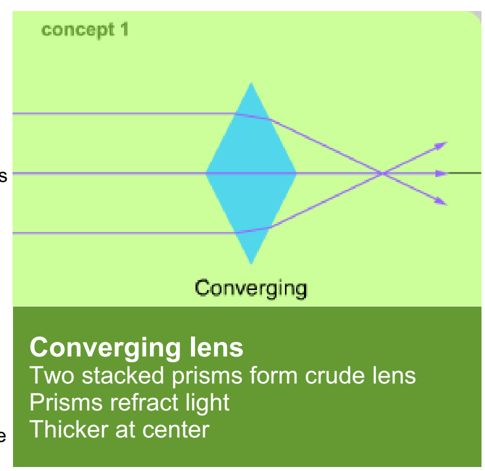
37.1 - Lenses

Lens: A device that uses refraction to redirect and focus light.

Lenses are used in eyeglasses to focus images.

Light changes direction − refracts − as it passes from air to a lens material and back to air. Lenses take advantage of refraction to create images.

In general, lenses are made of a transparent material with a curved surface or surfaces. Although lenses are typically curved, we will use two prisms, as shown to the right, to introduce the fundamental functioning of lenses.



In the upper diagram, you see two prisms placed so that their flat bases are touching. Three parallel rays of light strike the prisms. Because of the sloping surfaces of the prisms, the rays are refracted in different directions.

As you can see, after passing through the prisms, the three rays converge on the far side. This first set of prisms serves as a model for a *converging lens*. Converging lenses are thickest at their centers. With a two-prism lens, all the incoming rays would not converge at one point like the three we have chosen to show, which is why prisms do not make good lenses: The image they create will be blurry. A curved surface will focus the rays more precisely.

In the second diagram, we invert the prisms, so they touch tip to tip. The rays diverge after passing through them; they will not intersect on the far side of the prisms. This configuration provides a model for a *diverging lens*. The image created by a diverging lens is a virtual one. By extending the paths of the rays that have passed through the prisms backwards, we could locate the virtual image on the same side of the lens as the object that creates it.

In this chapter, we focus on lenses that are both spherical and thin. A spherical lens has two surfaces, each of which is a section of a sphere. The curvatures of the two sides of a lens can differ. Saying it is "thin" means the distance across a lens is small relative to the distance to the object that is the source of the image. Thin lenses provide an accurate model for optical equipment ranging from contact lenses to telescopes.

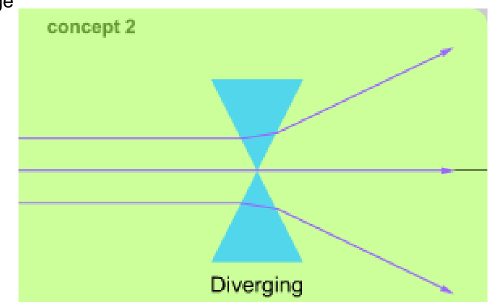
Much of the terminology and many basic concepts coincide for mirrors and lenses. A line that passes through the center of a lens perpendicular to its surface is called the *principal axis*. Parallel rays from an object infinitely far away refract through a lens and converge at the lens's *focal point*. The *focal length* f is the distance between the lens's center and its focal point. In contrast to mirrors, lenses have two focal points, one on each side of the lens, since incident rays can originate on either side of a lens. Rays that are close to the principal axis and converge at the focal point are called *paraxial rays*.

Lenses produce real and virtual images, just as mirrors do. As with mirrors, a real image can be projected onto a piece of paper, while a virtual image cannot. However, the locations of virtual and real images are reversed for lenses and mirrors. With lenses, virtual images occur on the **same** side of the lens as the object. Real images occur on the **opposite** side of the lens as the object.

The prisms in Concept 1 − the ones stacked base to base − produce a real image on the opposite side from the object. The ones placed tip to tip create a virtual image, one on the same side as the object.

As with mirrors, incoming rays that are far from the principal axis of a lens do not converge precisely at the focal point. This blurs the image and is called spherical aberration.

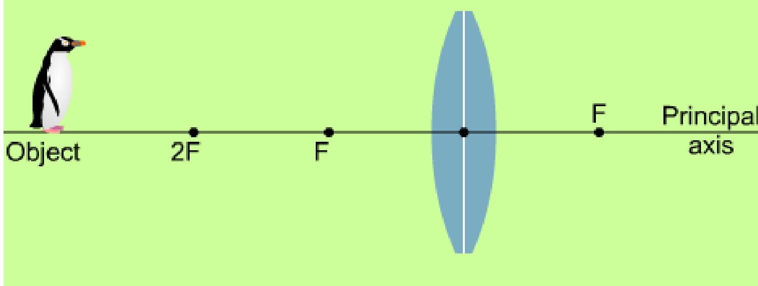
concept 2



Diverging lens
Thinner in center, light “spreads out”

37.2 - Converging lens: ray-tracing diagram

concept 1



Rays for a converging lens
What will be the nature of the image?

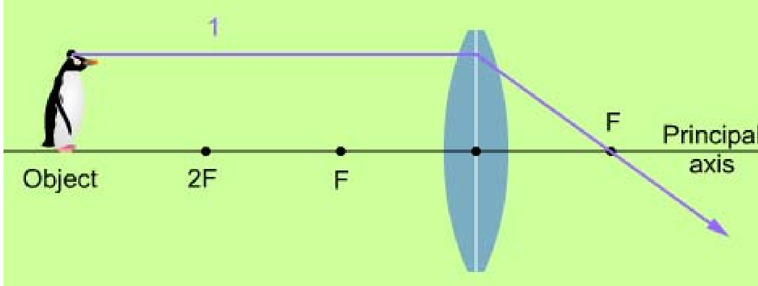
As with mirrors, ray-tracing diagrams are used to determine the orientation, and approximate size and location, of an image produced by a lens. The diagrams in this textbook employ a convention to make them easier to understand: Objects are shown fairly close to lenses that are relatively thick compared to the distances shown. For the approximations in the theory of thin lenses to hold true, the objects should be quite far from the lenses and the lenses should be much thinner than shown. Being literal would make for illustrations that would stretch far across the page or computer screen, so the diagrams are not drawn to scale.

Light that strikes the lens at an angle refracts twice, once upon entering the lens, and again while exiting. However, with a thin lens, we simplify things by modeling the light ray as changing direction just once as it passes through the lens.

Physicists use three specific rays to analyze the behavior of a lens; these rays differ somewhat from those used to analyze mirrors. A single converging lens can create different image types based on the position of the object relative to the focal point. In this example, we discuss the image created by an object more than twice the focal length away from a converging lens.

Ray 1 begins as a horizontal line, parallel to the principal axis, starting at the top of the object. It refracts at the lens and passes through the focal point on the far side of the lens. (We use “near” to describe the object’s side of the lens, and “far” to describe the other side. We will typically place objects on the left side of the lens. Images can be created on the near or far side of the lens.)

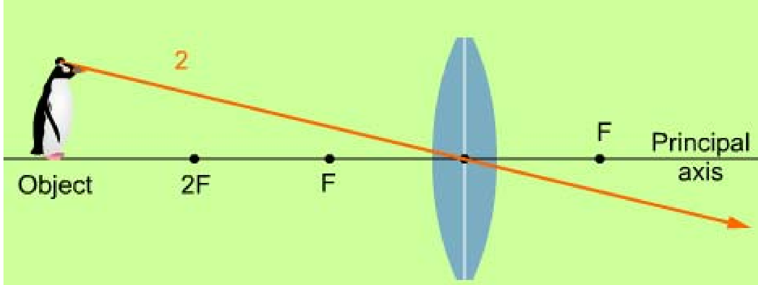
concept 2



Ray 1:
Parallel to axis, then passes through far focal point

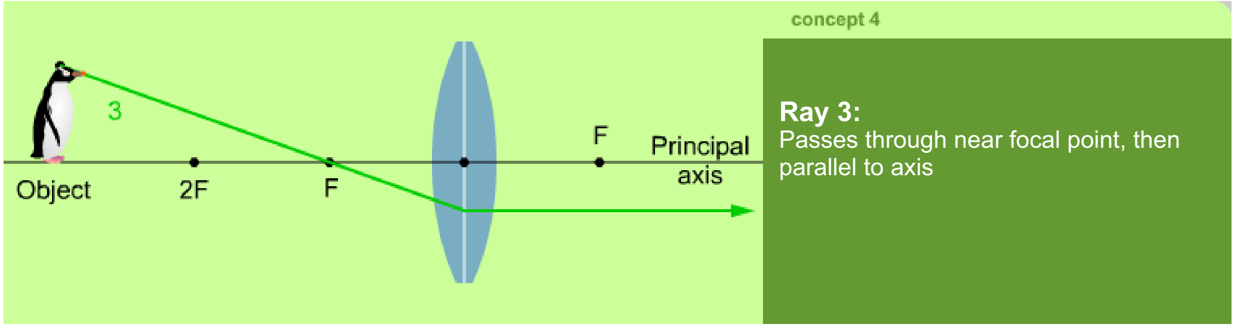
Ray 2 starts at the top of the object and passes through the center of the lens. It does not change direction.

concept 3

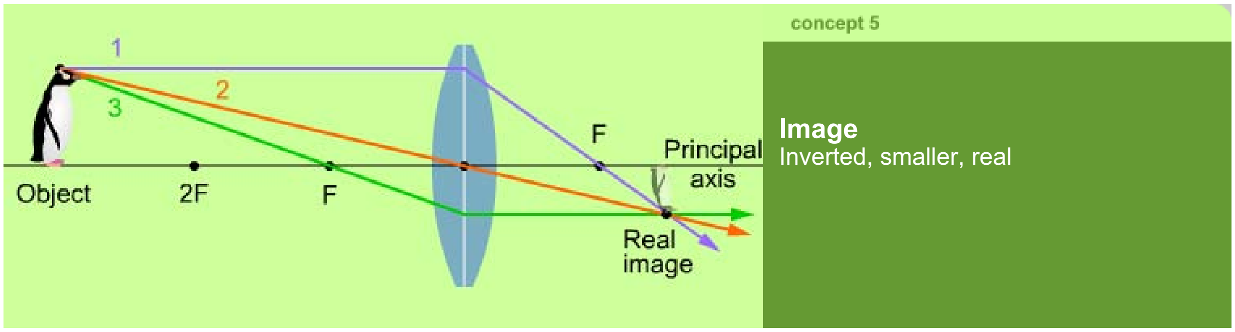


Ray 2:
Passes unchanged through center of lens

Ray 3 starts at the top of the object, passes through the focal point on the near side of the lens, is refracted by the lens, and continues parallel to the axis on the far side of the lens.

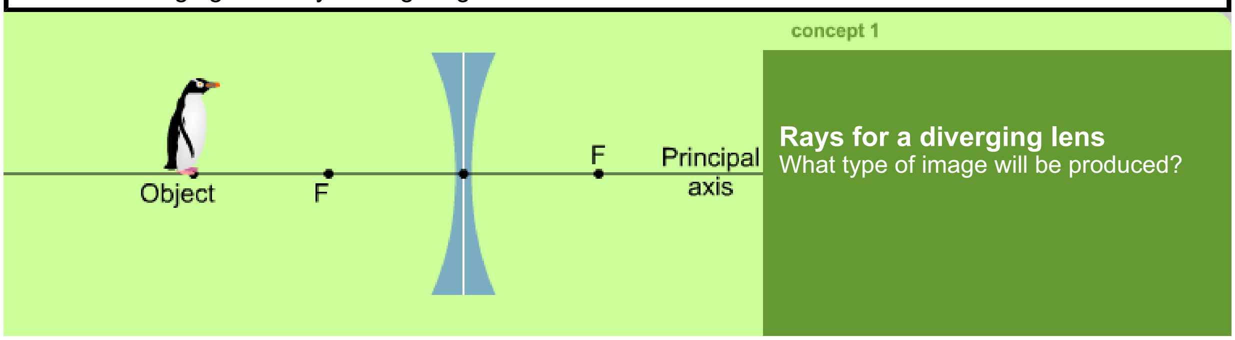


The image created by the lens is shown in the diagram below, where all three rays are combined together. The point at which the rays converge indicates the location of the top of the image. The image is inverted, smaller than the object and real.



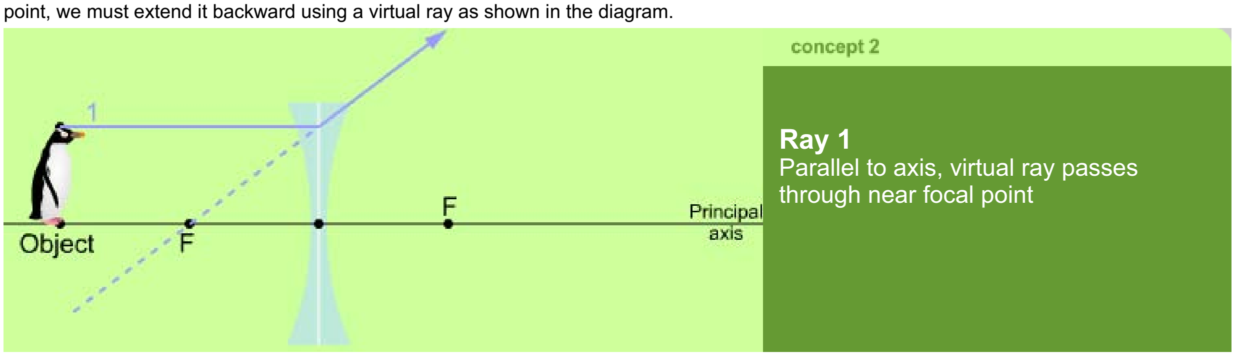
This example provides a model for a typical camera lens, which creates an inverted real image on the film or digital recording device on the far side of the lens.

37.3 - Diverging lens: ray-tracing diagram

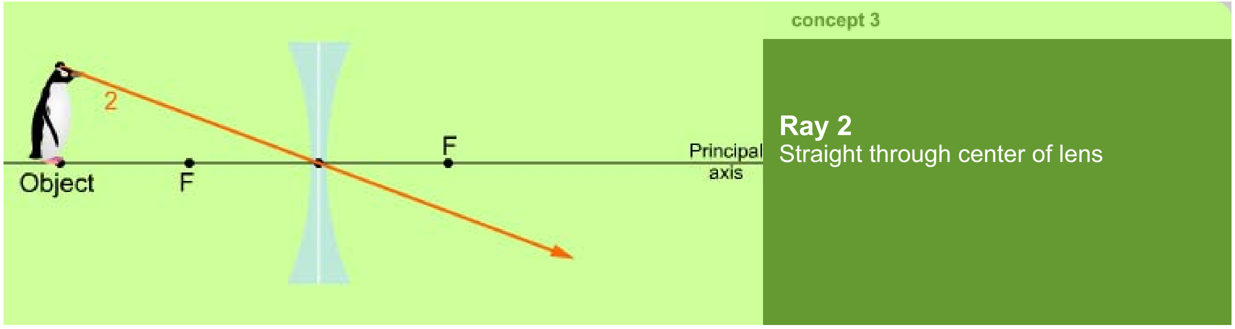


Like a convex mirror, a single diverging lens produces only one kind of image. The image is always upright, smaller than the object, and virtual. We use three rays to analyze the image.

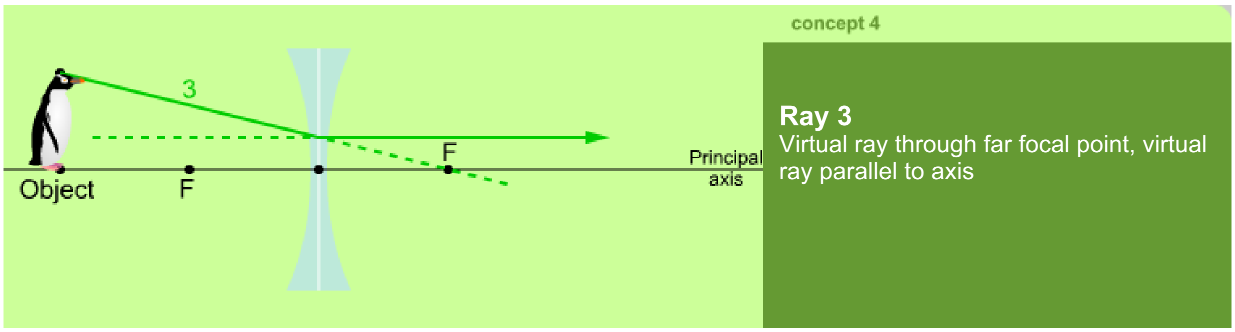
Ray 1 starts parallel to the principal axis. With a diverging lens, it refracts away from the principal axis. To have the ray pass through a focal point, we must extend it backward using a virtual ray as shown in the diagram.



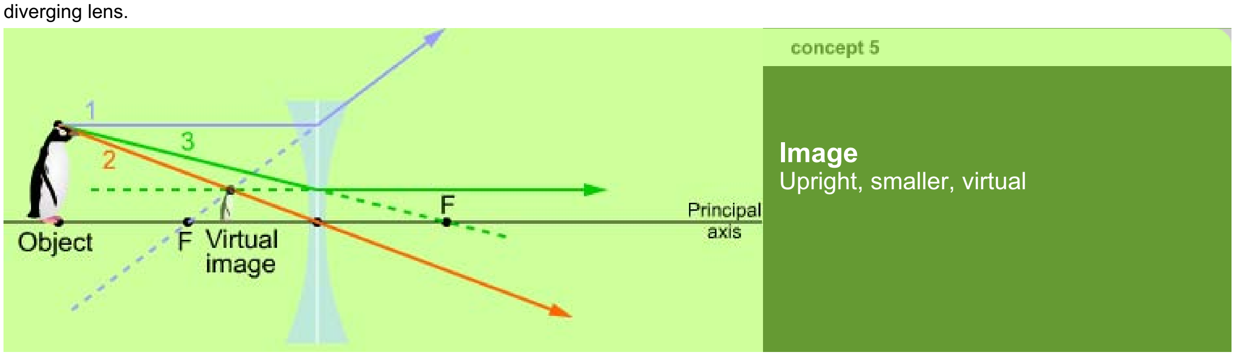
Ray 2 passes through the lens without changing direction.



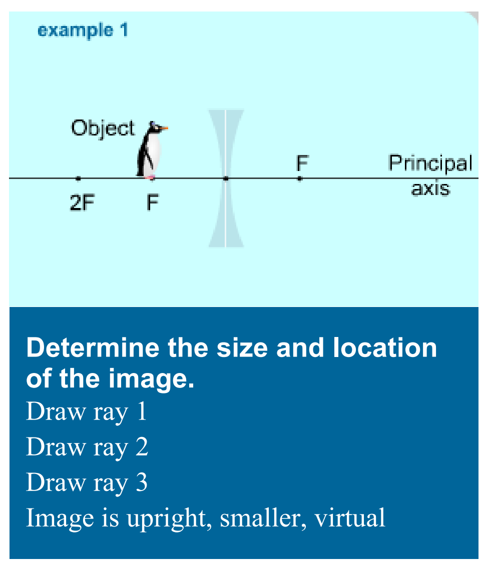
Ray 3 must refract to be horizontal on the far side of the lens. The forward extension of the incident portion of this ray is a virtual ray passing through the focal point on the far side. The backward extension of the refracted portion of the ray is a virtual ray parallel to the principal axis on the object side of the lens.



As the illustration below shows, the image is upright, smaller than the object, and virtual. This is the case for all images produced by a single diverging lens.



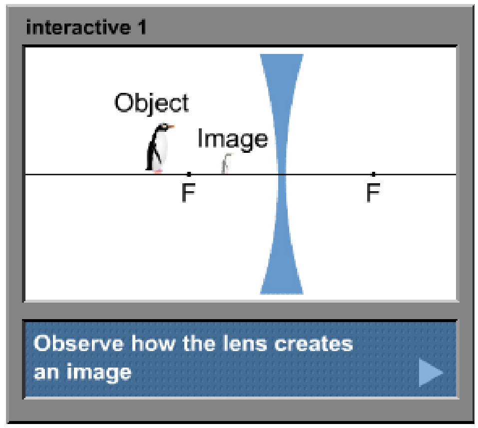
The example to the right leads you through another ray tracing exercise for a diverging lens. The object in the example is closer to the lens than the one above.



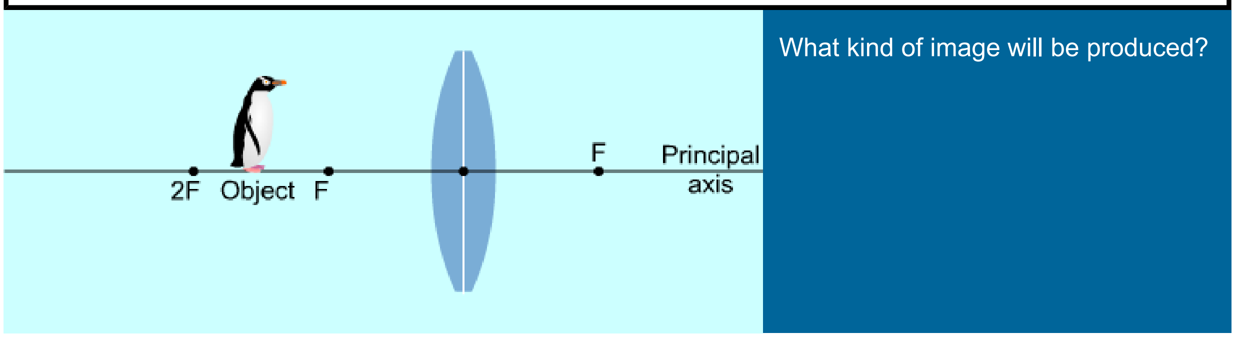
37.4 - Interactive problem: image with a diverging lens

This simulation lets you experiment with a diverging lens. You can drag the object toward and away from the lens, and note how the image changes. Also note which properties of the image remain consistent.

You can press “Show rays” to use the simulation to create a ray diagram. Pressing “Hide rays” turns off the rays.



37.5 - Sample problem: object outside the focal point of a converging lens

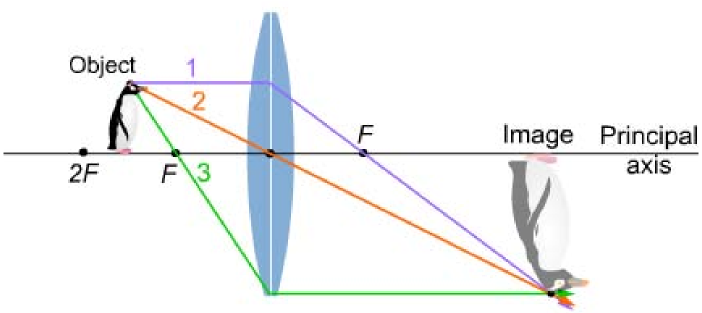


As shown above, the penguin is between the focal point and a distance twice the focal length. What kind of image will result? Once you know the result, can you think of any application that relies on a configuration like this?

Strategy

Use ray tracing to establish the nature of the image.

Diagram



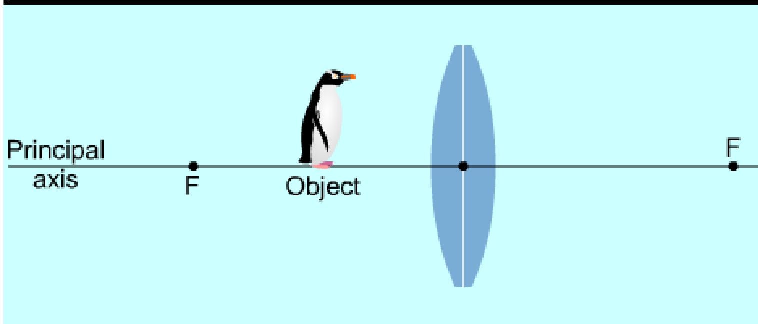
Step-by-step solution

1. Ray 1 starts parallel to the principal axis from the head of the penguin, refracts and passes through the focal point on the far side of the lens.
2. Ray 2 passes through the center of the lens without changing direction.
3. Ray 3 passes through the focal point on the near side and refracts at the lens to be parallel to the principal axis on the far side.

The result is an inverted image that is larger than the initial object. It is a real image, since light rays pass through the position of the image, and they would create a projected image on a screen placed there.

Above, we also asked: What might be an application that relies on a configuration like this? A movie theater projector is one example. The audience views enlarged, real images on the screen. Light is projected through the film, creating a luminous object from a tiny likeness, and then passes through a projection lens to be magnified. The real image appears on the movie screen. The “object” is upside down on the film, so when the lens inverts the image, it appears upright to the audience.

37.6 - Sample problem: object inside the focal point of a converging lens



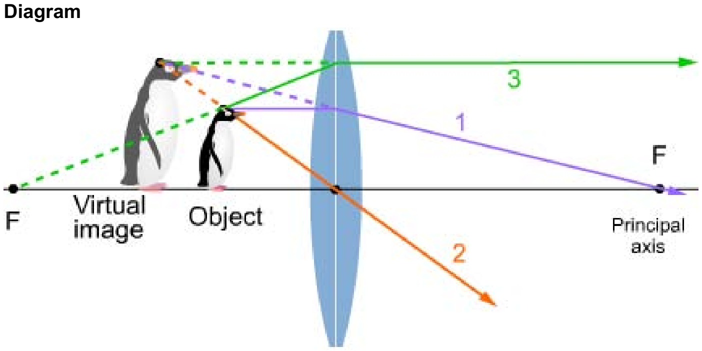
What kind of image will be produced?

The penguin is inside the focal point of a converging lens. What kind of image will result? Again, once you know the result, can you think of any application that relies on a configuration like this?

Strategy

Use ray tracing to establish the nature of the image.

Diagram



Step-by-step solution

1. Ray 1 starts parallel to the principal axis. It refracts at the lens and passes through the focal point on the far side. To obtain a convergence point for the image, we extend the ray backward, showing this extension with a dashed line.
2. Ray 2 passes through the center of the lens without changing direction. We again extend a virtual ray backward.
3. Ray 3 must be horizontal on the far side of the lens and it must also pass through a focal point. Here, it starts at the top of the object, directed upward so that its backward extension passes through the focal point on the near side. It refracts at the lens, becoming horizontal on the far side. By extending the horizontal portion backward, we create a virtual ray that passes through the convergence point.

The result is an upright, virtual image that is larger than the object. This configuration is typical of objects being viewed in a magnifying glass. The magnifying glass enlarges objects, making it easier to view their fine detail. The image cannot be projected, but it is upright, a matter of great convenience to physicists and textbook authors over 40 who may use a magnifying glass to see fine details!

37.7 - Lens equations

Quantity	Positive sign	Negative sign
Focal length, *f*	Converging lens	Diverging lens
Image distance, d_i	Far side (real)	Object side (virtual)
Object distance, d_o	Real	Virtual
Magnification, *m* and height, *h*	Image upright	Image inverted
Radius of curvature, *R*	Center on far side	Center on object (near) side

concept 1

Lens and magnification equations
Interpretation of signs

Ray diagrams provide a qualitative technique for establishing the nature of an image created by a thin lens. The image's position and size can be quantified with a set of three equations called the lens equations. These are the focus (we could not resist) of this section.

The *thin lens equation* is shown in Equation 1. It states the relationship between object distance, image distance, and focal length: The sum of the reciprocals of the object distance and the image distance equals the reciprocal of the focal length.

For instance, suppose you want to calculate the location of an image created by placing an object 0.50 meters away from a lens with a focal length of 0.33 meters. The thin lens equation and a little arithmetic reveal that the image distance is 1.0 meters. (We solve this problem in

Example 1.) The positive value for the image distance places it on the far side of the lens, which means the image is real. These values also confirm the answer to a question posed in an earlier section, where the nature of the image of an object between F and 2F was determined using ray tracing.

The formulas in Equation 2 deal with *lateral magnification*, which is defined in the same way for lenses as it is for mirrors: the ratio of the image height to object height. It can also be computed using distances: The lateral magnification equals the negative of the ratio of image distance to object distance. This is the second equation. For instance, in the example just discussed above, this ratio is −2.0: the image is twice as tall as the object. The negative sign means the image is inverted.

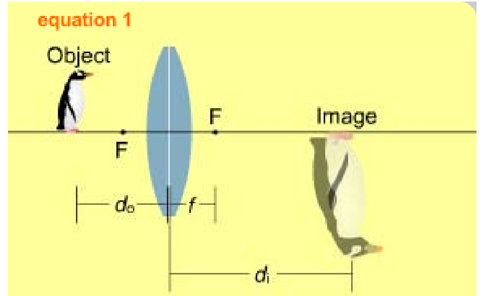
The equation in Equation 3 is the *lensmaker's equation*. It relates the focal length of the lens to some of its physical properties, specifically its index of refraction and the radius of curvature of each surface of the lens. The equation enables a lensmaker (or physics student) to calculate the focal length for a lens that has differing curvatures on the two sides of the lens, or to create a lens with a particular focal length.

We show the version of the equation for a lens in air. If the lens is immersed in another substance, say water, then the variable n has to be replaced by the ratio of the index of refraction of the lens to the index of the surrounding material (for example, $n = n_{\text{lens}}/n_{\text{water}}$). The object side of a microscope objective lens is often immersed in a special optical oil, as with the German "Oel Immersion" objective lens shown below.

The equations can be tricky to apply because various values like image distance can be either positive or negative. You need to use the table in Concept 1 above, or perform the more difficult task of memorizing the conventions.

To review perhaps the trickiest sign convention, the radius of curvature R in the lensmaker's equation is negative when the center of curvature is on the near (the object) side of the lens and positive when it is on the far side. If this sounds confusing, consider each lens surface as part of a sphere. If the sphere's center of curvature is on the object side, so the sphere would enclose an object relatively near the lens, then R is negative.

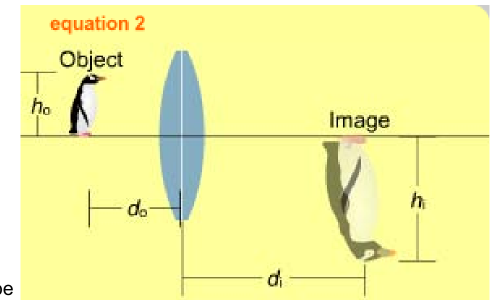
With a converging lens that is convex on both sides, like the ones used in the illustrations to the right, the radius of curvature for the near surface, R_1, is positive, and the radius of curvature for the far surface, R_2, is negative. As the diagram illustrates, with a lens that is convex on both sides, the center of curvature for each surface is on the opposite side of the lens from the surface. With a diverging lens made up of two concave surfaces, the signs are reversed and the centers are on the same side as the lens surfaces. Remember, we said this was tricky!



Thin lens equation

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

d_o = object distance
d_i = image distance
f = focal length



Magnification equation

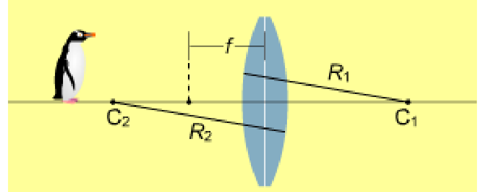
$$m = \frac{h_i}{h_o}$$

m = magnification
h_i = height of image
h_o = height of object

$$m = -\frac{d_i}{d_o}$$

There are additional conventions: Focal lengths are positive for converging lenses and negative for diverging lenses. Object distances are positive when the object is real and negative when it is virtual. (Virtual objects can arise when there are multiple lenses. A virtual object is on the side of a lens opposite to the source of the light.) Image distances are positive when the image is real and negative when it is virtual. Magnification is positive for an image that is upright relative to the object and negative for one that is inverted relative to the object.

equation 3

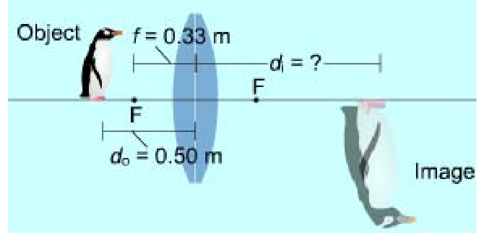


Lensmaker's equation, air

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

n = index of refraction
R_1 = radius of near surface
R_2 = radius of far surface

example 1



Find the distance to the image.

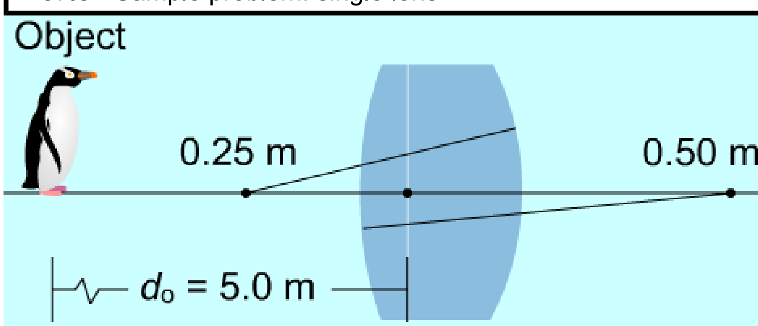
$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}$$

$$\frac{1}{d_i} = \frac{1}{0.33\text{ m}} - \frac{1}{0.50\text{ m}} = (3.0 - 2.0)$$

$d_i = 1.0$ m

37.8 - Sample problem: single lens



The index of refraction of the lens is 1.50. What will be the location of the image?

Variables

index of refraction of glass	$n = 1.50$
radius of near surface of lens	R_1
radius of far surface of lens	R_2
focal length of lens	f
distance of object	$d_o = 5.0$ m
distance of image	d_i

Strategy

1. Apply the lensmaker's equation to the index of refraction and the radii of curvature to determine the focal length.
2. Use the focal length just calculated, along with the stated object distance, to determine the location of the image using the thin lens equation.
3. In all your work, be careful about signs, and make sure not to confuse R_1 and R_2.

Physics principles and equations

We will use the lensmaker's equation and the thin lens equation.

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

In the lensmaker's equation, R_1 is the radius of curvature for the **near** surface of the lens. The value for a radius is negative when the center of curvature is on the object side.

Step-by-step solution

First we use the lensmaker's equation to find the focal length of the lens.

Step	Reason
1. $R_1 = +0.50$ m	Near surface: center of curvature on far side
2. $R_2 = -0.25$ m	Far surface: center of curvature on near side
3. $\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$	lensmaker's equation
4. $\frac{1}{f} = (1.50-1)\left(\frac{1}{+0.50\text{ m}} - \frac{1}{-0.25\text{ m}}\right)$ $\frac{1}{f} = 0.50\,(2.0+4.0)$ $\frac{1}{f} = 3.0\text{ m}^{-1}$	substitute; evaluate $1/f$

Now we use the reciprocal of the focal length calculated above together with the given object distance to find the image distance using the thin lens equation.

Step	Reason
5. $\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$	thin lens equation
6. $\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}$	rearrange
7. $\frac{1}{d_i} = 3.0\ \text{m}^{-1} - \frac{1}{5.0\ \text{m}}$ $\frac{1}{d_i} = 2.8\ \text{m}^{-1}$	substitute; evaluate $1/d_i$
8. $d_i = 0.36\ \text{m}$	solve for d_i

The image distance is 0.36 meters. It is a real image. Since the image distance is positive, it is on the far side of the lens.

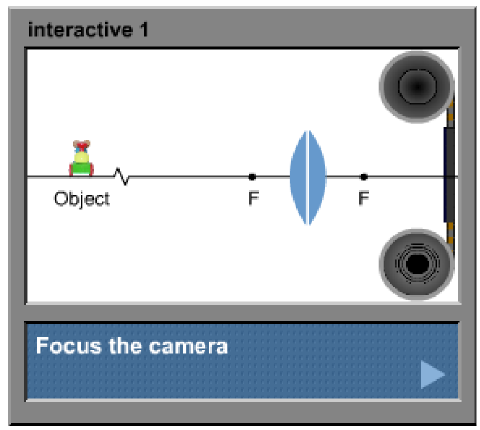
37.9 - Interactive problem: focus a camera

You are taking a picture of the driving mouse shown to the right. You must set the lens distance from the film in order to take a properly focused picture.

The mouse is 5.00 meters from the center of the lens, significantly farther from the lens than the film, a fact we show with the zig-zag in the principal axis line. The focal length of the lens is 4.00 cm. Moving the lens does not significantly change the distance to the mouse, so you can treat the 5.00-meter distance as constant.

Set the distance between the lens and the film (where the image is projected) in the simulation to the nearest 0.01 cm, based on the calculations you make. You can grab the lens to position it roughly in the right place, and then fine-tune the position with the controls below. When you are ready, press CLICK to take the picture. If you have calculated correctly, you will see a sharp image.

If you find this problem difficult, consult the section with the thin lens equation.

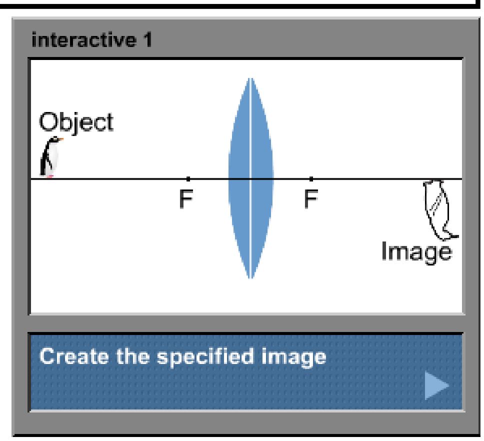


37.10 - Interactive problem: optical bench with a lens

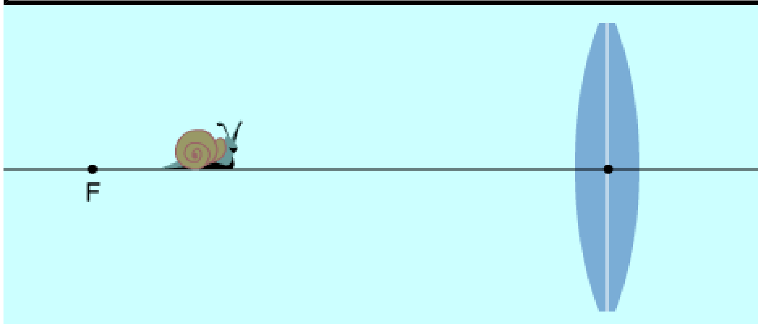
Your mission here is to create the image at the location shown in the graphic at the right. The image is 13.5 cm tall, and 15.0 cm from the lens. The object is 9.0 cm tall. (Note: We are deliberately being ambiguous about mathematical signs here!)

You set both the object location and the lens's focal length to the nearest 0.01 cm in this simulation. Press CHECK to test your answer.

A hint: Start this problem by determining the magnification required for the specified image height, and then consider how that helps you to specify the object distance. If you need help with this problem, consult the section on lens equations. You will need to apply three equations from that section.



37.11 - Interactive checkpoint: lens equations



A 1.80 cm tall snail sits 9.00 cm in front of a thin lens with focal length 11.5 cm. Find the image distance, the image height, and the lateral magnification of the image.

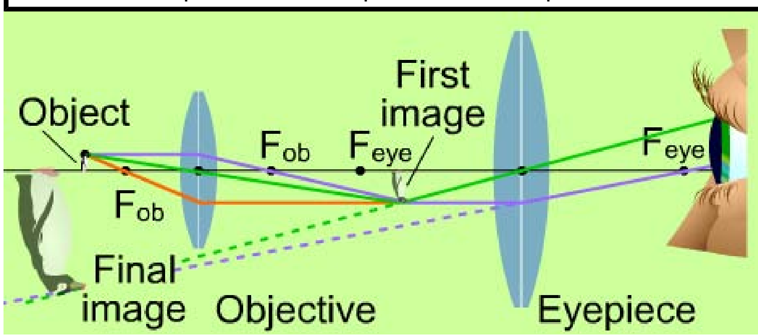
Answer:

$d_i =$ ______ cm

$h_i =$ ______ cm

$m =$ ______

37.12 - Multiple lenses: compound microscope



concept 1

Compound microscope

Two lenses – objective and eyepiece

Objective focal length very short

First image real, near eyepiece focal point

Final image inverted, magnified, virtual

Multiple lenses underlie the design of many microscopes, telescopes and other pieces of optical equipment. Ray-tracing diagrams and the equations for lenses can be applied to analyze configurations that involve multiple lenses. We use a ray-tracing diagram in this section to explain the functioning of a *compound microscope*. You see a conceptual diagram above of this two-lens system. The diagram is not drawn to scale.

The first lens the light passes through, called the *objective*, has an extremely short focal length (this length is exaggerated in the diagram above). It creates the inverted real image shown. This image serves as the object for the second lens, the *eyepiece*, which is the lens nearest to the eye. The eyepiece typically has a focal length of a few centimeters. A distance much greater than the focal length of either lens separates them.

In the diagram, the object is not far outside the focal point of the objective lens. This creates a real image that is at or just inside the focal point of the eyepiece. The eyepiece then creates an inverted virtual image that is much larger than the initial object.

The compound microscope is designed not only to produce a greatly magnified image, but with two other goals as well. One is to keep the instrument relatively compact. This is accomplished by using an objective lens with a very short focal length so that the object can be very close to it and yet remain outside that point.

The eyepiece also creates a final virtual image that is far away from the viewer. It is designed this way because the human eye is most relaxed when viewing distant objects.

37.13 - The human eye

The human eye contains a variable lens. This organ − a remarkable product of evolution − employs the lens to focus images on the *retina* at its back surface. Special cells called rods and cones line the retina. Light stimulates these cells, and they send signals in the form of electric impulses to the brain via the optic nerve.

Light passes through four different components of the eye on its way to the retina: the cornea, the aqueous humor, the lens, and the vitreous humor. Each has a different index of refraction, ranging from about 1.30 to 1.40. This means the greatest amount of refraction occurs when the light first crosses from the air $(n = 1.00)$ to the cornea, where the indices differ the most.

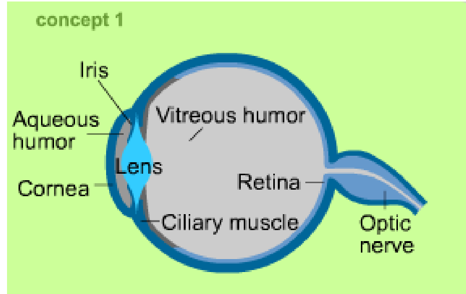
In order for an object to be perceived clearly, its image must be focused on the retina, at a fixed distance from the lens. The real image projected on the retina is upside down, but the brain automatically corrects for this.

The eye must be able to create crisp images of objects that are located at varying distances. It accomplishes this by changing the shape of the lens, a process called *accommodation*. By causing the lens to contract or expand, the eye changes the radius of curvature and the focal length of the lens. On the right, you see the lens in both a relaxed state for viewing objects far away, and tensed as its curvature has been changed to focus on a nearer object.

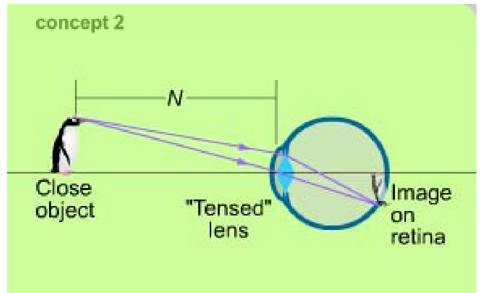
The closest distance at which an object can be brought into focus is called the *near point*. A young adult with a fully tensed lens can clearly see objects that in some cases are as close as 15 cm, with an average near point for this age being 25 cm. By the time a person reaches her early forties, the near-point distance increases to an average of 40 cm. By age 65, the average near-point distance is 400 cm. Changes in the eye that occur with aging explain this increase. In fact, this progression is so predictable that the distance to the near point can pinpoint age to within a few years.

The *far point* is the greatest distance at which an object can be seen clearly. For people who do not need glasses for distance vision, this point is effectively infinitely far away. For those who are nearsighted, it can be much closer.

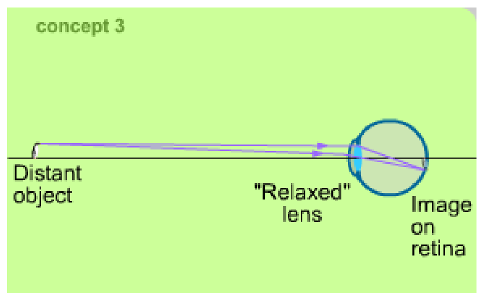
If you have 20/20 vision, you are lucky: You perceive objects crisply without the use of prescription lenses. If you have 20/60 vision, it means that you must be as close as 20 feet to see what a person with typical vision can see at 60 feet. Baseball players are noted for having acute vision, such as 20/15 vision. This means they see at 20 feet what most people need to be within 15 feet to see clearly.



Human Eye
Contains variable lens
Image forms on retina

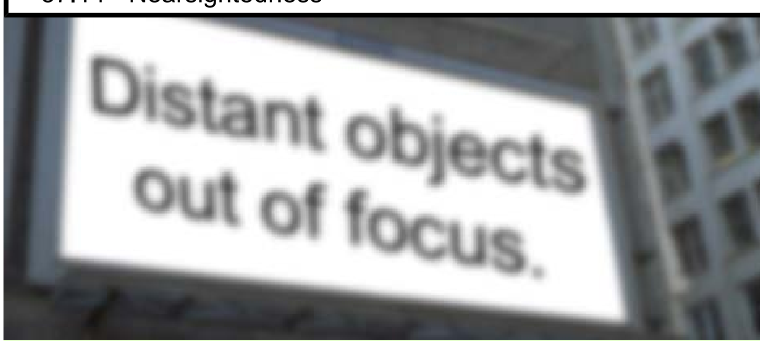


Near point
Closest in-focus object distance



Far point
Farthest in-focus object distance

37.14 - Nearsightedness



concept 1

Nearsightedness
Distant objects out of focus

Nearsightedness: Ability to focus on nearby objects but not distant objects.

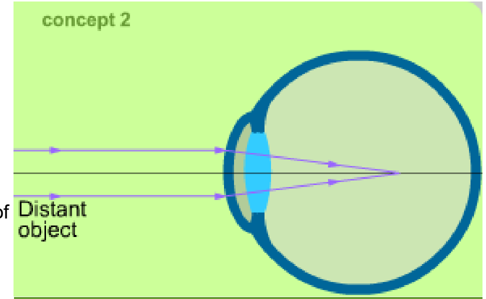
To people with nearsightedness, or *myopia*, faraway objects are blurry and look like the image above.

To view an object that is far away, the eye's ciliary muscle completely relaxes, causing the lens to be as flat as possible. In a nearsighted eye, the eye is too long, causing images of faraway objects to form in front of the retina, even when the lens is completely relaxed. You see this in the conceptual diagram in Concept 2.

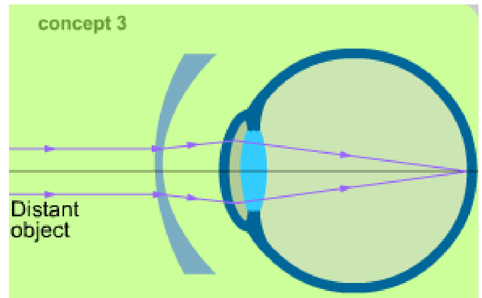
To correct this problem in focusing, a diverging lens can be placed in front of the eye. As illustrated in Concept 3, the lens spreads out the parallel rays before they reach the eye, so that the rays do not come to a focus quite so soon, but farther back upon the retina where they should.

More recently, laser surgery has become a popular method for correcting nearsightedness. A laser directed at the eye reshapes the cornea, causing its outer surface to be flatter so that an image can form more crisply on the retina. The angle of incidence for incoming light rays is less, causing them to converge farther back in the eye, at the retina.

Contact lenses can also correct nearsightedness; they change the radius of curvature of the front surface of the eye. Since the contact lenses are thicker at the edges than the center, they work by flattening the eye's surface.

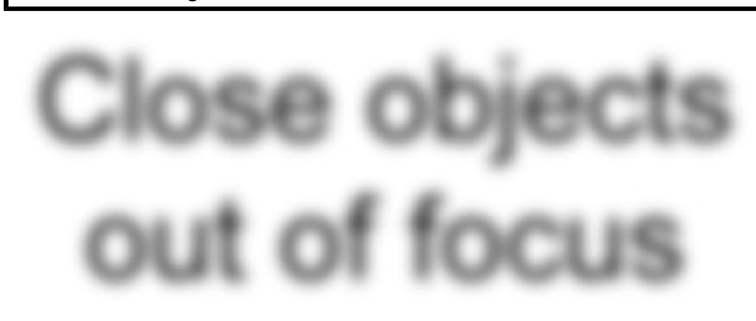


The problem
Image forms in front of retina



A solution
Corrected with diverging lens

37.15 - Farsightedness



concept 1

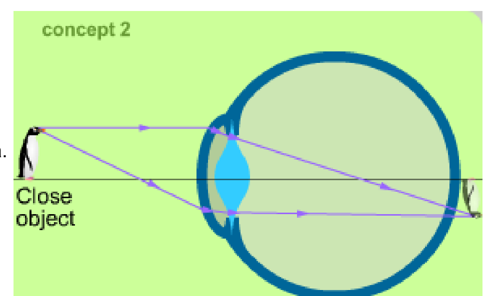
Farsightedness
Close objects out of focus

Farsightedness: Inability to clearly see objects that are relatively close.

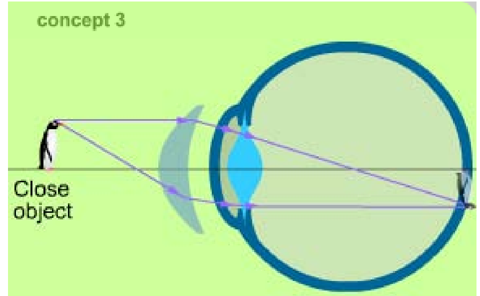
With farsightedness, or *hyperopia*, the image of a nearby object forms behind the retina. The eyeball is too short, and no matter how hard the ciliary muscle strains, it cannot contract the lens enough to focus the image on the retina. You see this shown in Concept 2.

People with hyperopia have difficulty reading things that are close and often use reading glasses as a result. A related condition called *presbyopia* ("old vision") occurs in many people starting in their forties. With presbyopia, the eyeball is not too short, but the lens itself has become stiff with age and is unable to contract enough to focus images on the retina.

A converging lens can compensate for this problem by creating a virtual image beyond the eye's near point, so that the lens can focus an image on the retina. This solution is shown in Concept 3.



The problem
Image forms behind retina



A solution
Corrected with converging lens

37.16 - Physics at work: laser eye surgery

Laser eye surgery, a procedure now quite popular, provides a way to alter the shape of the cornea so that glasses or contact lenses are not needed, or are required less frequently. It has proven quite popular amongst both celebrities (such as Tiger Woods) and non-celebrities.

To explain how the procedure works, we first review some fundamentals of how the eye focuses an image. The eye has two components that are most responsible for the location of the image it creates: the cornea (its outermost surface) and an internal variable lens. Light refracts as it passes through both of these components.

Although one might think most of the refraction occurs at the lens, in fact about two-thirds of the refraction occurs at the cornea. It occurs there because the index of refraction of the cornea differs substantially from that of air. This difference is much larger than differences in the indices of refraction within the eye. The cornea and the lens combine to focus the light at the retina, which is at the back of the eye. The eye adjusts the curvature of the lens in order to alter its focal length, enabling it to focus objects at various distances. When the eye can change the lens shape enough so that light converges crisply at the retina for objects both near and far, a person has no need for glasses.

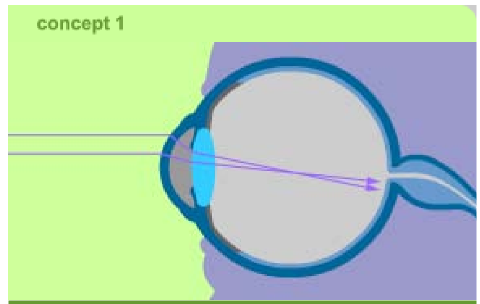
However, many people need glasses or contact lenses in order to see clearly. Nearsightedness and farsightedness occur when the lens's shape cannot be changed enough to sharply focus images on the retina. Eyeglasses and contact lenses compensate for this by providing a second lens that assists in the process of focusing the image at the proper position.

Laser eye surgery takes a different tack: It changes the eye itself. The surgery changes the shape of the cornea in order to correct for the limitations of the lens. The procedure does not alter the lens inside the eye; rather, it "tunes" the cornea.

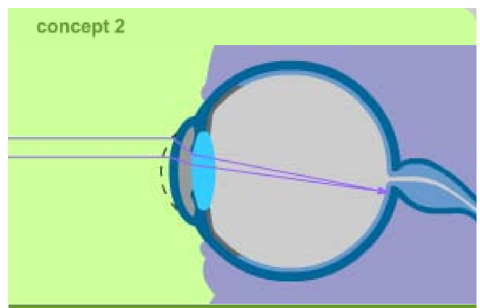
Currently, a type of laser surgery called *LASIK* (*laser in situ keratomileusis*) is the most widely used. To begin the process, medical personnel first determine how the cornea's curvature must be modified in order to improve vision. Then, a surgeon temporarily peels back the epithelium (the thin outermost layer of the eye) and trains a laser on the eye to remove small amounts of the cornea. The epithelium is then put back to cover the eye's surface.

Laser surgery can be used to treat both nearsightedness and farsightedness. Nearsightedness occurs when the lens cannot become sufficiently flat, resulting in an image formed in front of the retina. This is shown in Concept 1. To compensate for this, the surgeon flattens the cornea. This reduces the angle of incidence for light entering it, and the angle of refraction as a result. Concept 2 shows the eye after laser eye surgery for nearsightedness. The dashed line shows the original curvature of the eye.

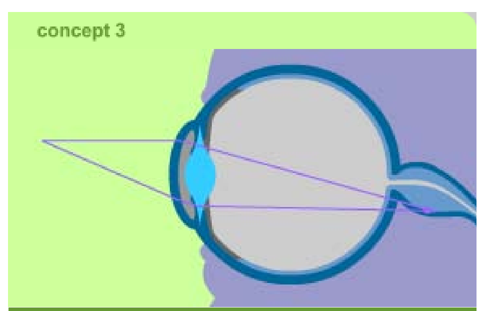
With farsightedness, the muscles of the eye cannot contract the lens enough. Close objects are focused behind the retina, again causing blurring. This problem is shown in Concept 3. To address it, the surgeon increases the steepness of the cornea. Since the angle of incidence of incoming light rays is increased, they refract more. The bottom illustration shows a farsighted eye after laser eye surgery. Again, the dashed line represents the initial shape of the eye.



The problem: nearsightedness
Image forms in front of retina

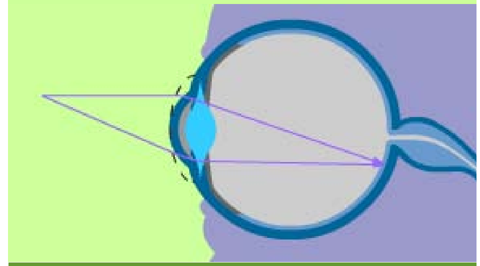


The solution
Cornea flattened to correct vision
After surgery – image forms on retina



The problem: farsightedness
Image forms behind retina

concept 4



The solution
Cornea steepened to correct vision
After surgery – image forms on retina

37.17 - Refractive power: diopters

Off-the-shelf reading glasses rated in diopters.

concept 1

Refractive Power
Inverse of focal length
Measured in diopters

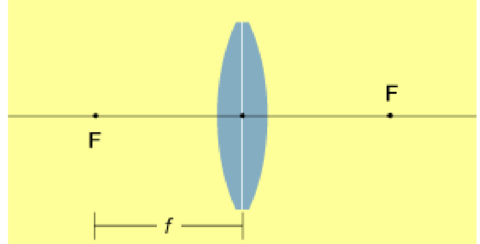
Refractive power: The reciprocal of the focal length. Often used by opticians and optometrists, who specify it in diopters.

If you wear glasses or contact lenses, you may have seen the term *diopters* in your prescription or on a contact lens box. The generic corrective reading glasses sold in drugstores are rated in diopters as well. The diopter is the unit for refractive power. Refractive power is the reciprocal of the focal length measured in meters.

A lens with greater refractive power refracts light more and has a shorter focal length than one with less power. Glasses and contact lenses for nearsightedness can have refractive powers as large as −20 diopters. Those for farsightedness typically range as high as +14 diopters.

The reason for the negative diopters is that myopia is corrected with diverging lenses, and the focal lengths of such lenses are negative. The drugstore reading glasses shown above correct for farsightedness.

equation 1



Refractive Power

$$P = \frac{1}{f}$$

P = refractive power in diopters (m^{-1})
f = focal length in meters (m)

example 1

What is the refractive power of these eyeglasses?

$f = 0.25$ m

$P = 1/f$

$P = 1/(0.25\text{ m})$

$P = 4.0$ diopters

37.18 - Angular size

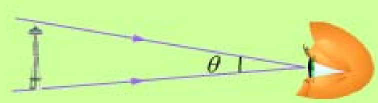
Angular size: How much of the field of vision is filled by an object or image.

Sizing up the Needle.

Imagine yourself in Seattle, standing two kilometers from the Space Needle. You look at the structure, and then raise your hand so that your thumb is a half-meter from your eyes. Your thumb now appears to be about the same size in your field of vision as the Space Needle, although you know that the Space Needle is much bigger.

This effect occurs because of the relative angular sizes of your thumb and the Space Needle. The angular size refers to the angle, measured in radians, *subtended* by an object when viewed from a particular distance. The subtended angle is perhaps as well explained with a diagram as in text. The subtended angle is labeled θ in Concept 1 to the right. It measures the angle of your vision "blocked out" by the object.

concept 1

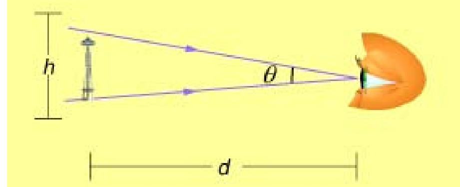


Angular size
Amount of visual field filled by object
Measured as angle

Trigonometry can be used to determine angular size when the height h of the object and the distance d to the object are known. The angular size can be approximated as the height divided by the distance to the object. This approximation uses a small-angle approximation, $\theta \approx \tan\theta$ for small angles, where the angle must be measured in radians. This approximation for angular size is very good (within one percent) for angular sizes less than 0.17 rad (about $10°$). Approximations like this can be useful in astronomy to approximate distance or size. For instance, it can be useful to know that the Moon subtends about $0.5°$ (0.009 radians). If you know its distance, you can quickly approximate its diameter using this fact and some trigonometry.

To leave space and return to Seattle: The thumb in the photograph above is 5 cm or so tall and 60 cm away. It has an angular size of 5 cm divided by 60 cm, which equals 0.08 rad. At a distance of 2 km, the Space Needle, which is 184 meters tall, would subtend an angle of about 184 m/ 2000 m, or 0.09 rad. These values confirm that the thumb is about the same angular size as the Space Needle.

equation 1

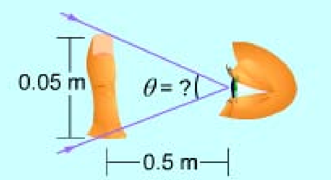


Angular size

$$\theta \approx \frac{h}{d}$$

θ = angular size (in radians)
h = height of object
d = distance to object

example 1



What is the angular size of this person's thumb at the given distance?

$\theta_{\text{Thumb}} \approx 0.05\text{ m} / 0.5\text{ m}$

$\theta_{\text{Thumb}} \approx 0.1\text{ rad}$

example 2

How far is this person from the Space Needle? She measures the angular size of her thumb and of the Needle as 0.11 radians.

$\theta_{\text{Needle}} \approx h_{\text{Needle}}/d_{\text{Needle}}$

$d_{\text{Needle}} \approx h_{\text{Needle}}/\theta_{\text{Needle}}$

$d_{\text{Needle}} \approx 184\text{ m}/0.11$

$d_{\text{Needle}} \approx 1700\text{ m}$

37.19 - Angular magnifying power: simple magnifier

In this section, we will use the concept of angular size to show how to calculate the angular magnifying power provided by a magnifying glass.

We start by asking you to consider the process of examining an object with fine detail. At first, you bring it as close to your eye's near point as you can. At the near point, it subtends an angle at your eye represented by θ_0, which is the largest it can be without getting out of focus. You see this illustrated in Concept 1.

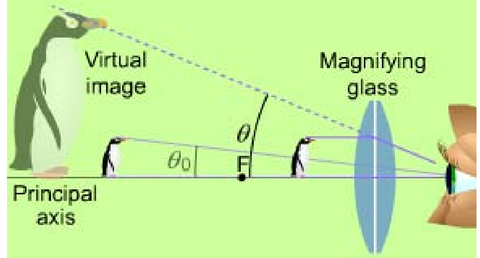
If you still cannot see the detail you want, you may fetch a magnifying glass and place it close to your eye. You bring the object closer to your eye than before, at a point that is also inside the focal point of the magnifying glass. At this point it forms a virtual image, which subtends an angle θ.

The *angular magnifying power* of the lens is defined as θ/θ_0. This is the ratio of the angle subtended by the image to that subtended by the object without the lens. You see the definition of angular magnifying power in Equation 1. The symbol for angular magnifying power is M. Be careful not to confuse this with the lateral magnification m of an image.

In Equation 2 we apply angular magnifying power to a magnifying glass. For a magnifying glass, we place the object at the near point to determine θ_0, since at this point its angular size is as large as possible while remaining in focus, providing a good measure of this instrument. For other instruments you will encounter later, such as a refracting telescope, we use a different object position to determine θ_0. (For a telescope, you would not place the object at the near point to determine its "unmagnified" angular size, since the object is likely to be very far away.)

In Equation 3 we show two particular cases of the magnifying glass equation. The first is when the image is created at the near point of the eye. At this point, the image has the maximum angular size possible while remaining in focus. The other case of interest is when the image is created at infinity. This configuration allows the eye to be completely relaxed when viewing the image, and is of special interest to jewelers and others who spend large amounts of time looking through a type of magnifying glass called a *loupe*. The magnification in this case is less than when the image is at the near point, but this can be compensated for by reducing the focal length of the lens.

concept 1



Angular magnifying power
Ratio of angular sizes of image, object

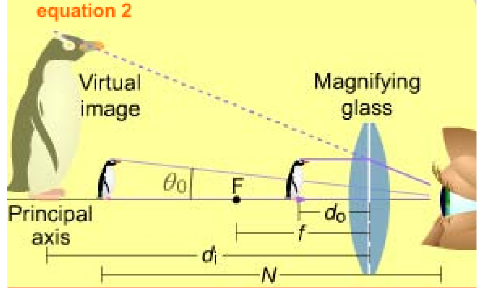
equation 1

Angular magnifying power

$$M = \frac{\theta}{\theta_0}$$

M = angular magnifying power
θ = image angular size
θ_0 = object angular size

equation 2



Magnifying glass

$$M \approx \frac{N}{d_o} = \left(\frac{1}{f} - \frac{1}{d_i}\right)N$$

N = near point distance
d_o = object distance
f = focal length
d_i = image distance

equation 3

Two cases of the magnifying glass

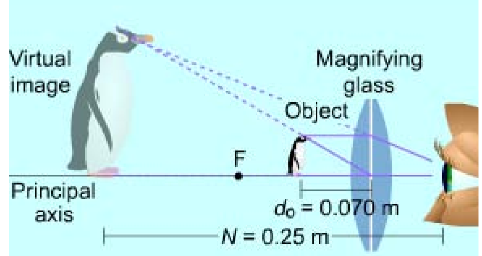
Image at near point: $d_i \approx -N$

$$M \approx \frac{N}{f} + 1$$

Image at infinity: $d_i = -\infty$

$$M \approx \frac{N}{f}$$

example 1



The image is at this person's near point, 0.25 m. The object is 0.070 m away. What is the angular magnifying power of the magnifying glass?

$M = N/d_o$

$M = 0.25\text{ m} / 0.070\text{ m}$

$M = 3.6$

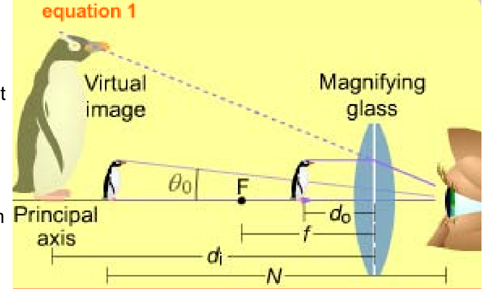
37.20 - Derivation: simple magnifier equation

The equation for the angular magnifying power of a simple magnifier, or magnifying glass, is derived below.

We start with the definition of angular magnifying power: the angular size of the image divided by that of the object. The angular size of the object is determined by placing it at the near point of the eye, since at this point it subtends as large an angle as possible while remaining in focus.

When determining the image angular size, we assume that the eye is relatively close to the lens, so that the distance to the image is approximately the same as measured from the eye or the lens. We also assume that since objects viewed with a magnifying glass are small compared to your distance from them, we can use the small angle approximation when determining the angular size.

equation 1



Magnifying glass

$$M \approx \frac{N}{d_o} = \left(\frac{1}{f} - \frac{1}{d_i}\right)N$$

M = angular magnifying power

N = near point distance

d_o = object distance

f = focal length

d_i = image distance

Variables

angular magnifying power	M
angular naked-eye size of object at near point	θ_0
angular size of image	θ
near point distance	N
image height	h_i
image distance	d_i
object height	h_o
object distance	d_o
focal length of lens	f

Strategy

1. Start with the definition of angular magnifying power.
2. Use the geometry of the diagram and the small-angle approximation to express the angular magnifying power in terms of distances, rather than angles.
3. Use the equations for lateral magnification to simplify, then use the thin lens equation to express the angular magnifying power in terms of the focal length, image distance, and near point distance.

Physics principles and equations

Angular magnification

$$M = \frac{\theta}{\theta_0}$$

Lateral magnification

$$m = \frac{h_i}{h_o} = -\frac{d_i}{d_o}$$

Thin lens equation

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

Mathematics principles

The small-angle approximation states that for a small angle θ measured in radians,

$$\theta \approx \tan\theta$$

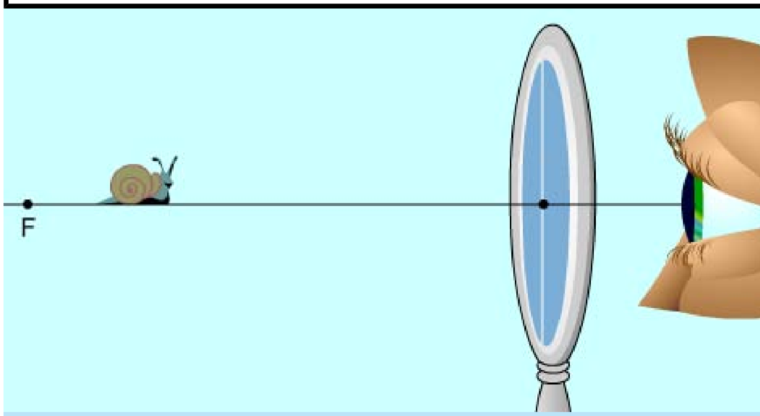
Step-by-step derivation

Step	Reason
1. $M = \frac{\theta}{\theta_0}$	angular magnifying power
2. $\tan\theta_0 = \frac{h_o}{N}$	inspection of diagram
3. $\tan\theta \approx -\frac{h_i}{d_i}$	inspection of diagram
4. $\theta_0 \approx \frac{h_o}{N}, \ \theta \approx -\frac{h_i}{d_i}$	small angle approximation
5. $M \approx \frac{-h_i/d_i}{h_o/N}$	substitute step 4 into equation 1

Now we use the lateral magnification equation and the thin lens equation to derive the general equation for a magnifying glass.

Step	Reason
6. $\frac{h_i}{h_o} = -\frac{d_i}{d_o}$ $-\frac{h_i}{d_i} = \frac{h_o}{d_o}$	lateral magnification
7. $M \approx \frac{h_o/d_o}{h_o/N} = \frac{N}{d_o}$	substitute equation 6 into equation 5 and simplify
8. $\frac{1}{d_o} = \frac{1}{f} - \frac{1}{d_i}$	thin lens equation
9. $M \approx \left(\frac{1}{f} - \frac{1}{d_i}\right)N$	substitute equation 8 into equation 7

37.21 - Interactive checkpoint: simple magnifier



Nolan uses a thin lens with focal length 11.5 cm as a magnifying glass to examine a snail, which is 1.80 cm tall. He adjusts the position of the lens so that he sees a focused image at (negative) infinity. The image has an angular magnification of 2.45. What is the approximate near point distance of his eye?

Answer:

$N \approx$ ______ cm

37.22 - Telescopes

concept 1

Refracting telescopes
Objective lens gathers light
Eyepiece magnifies image

Simple refracting (lens-based) telescopes have two lenses. A clever configuration of these lenses enables the telescope to serve three purposes: to gather as much light as possible from the object, to magnify the object and to provide as relaxed viewing of the image as possible.

The first lens encountered by incoming light rays, the objective lens, is used to gather as much light as possible. Light is at a premium in telescopes since faraway objects are typically faint. You cannot see most of the stars in the sky because they are too dim, not because they are too small (in angular size). Astronomers are willing to pay high prices to obtain telescopes with large objective lenses. Amateur astronomers purchasing their first telescope are often advised to pay more attention to its light gathering capability than to its magnifying power.

The eyepiece of a telescope performs two important tasks: It magnifies the real image created by the objective lens and creates a virtual image whose image distance is essentially at infinity. Creating a distant image is useful because the human eye is most relaxed when viewing objects at a distance.

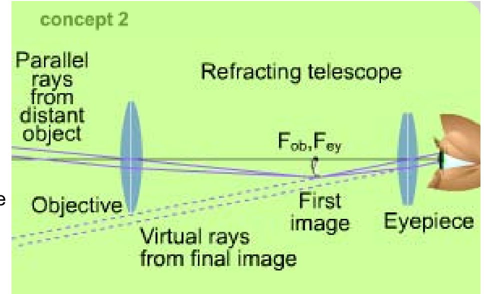
At the right, you see the combination of lenses that constitutes a basic refracting astronomical telescope. To analyze how this system functions, you must first determine the location of the real image created by the objective lens. This image serves as the object for the eyepiece.

Telescopes are pointed at faraway objects, so the light rays entering the telescope are essentially parallel. Both the objective lens and the eyepiece are converging lenses. The objective lens creates a real image of the distant object at its focal point. The lenses are arranged so that this point is also at the focal point on the incident side of the eyepiece. Putting the object here causes the virtual image created by the eyepiece to be at infinity. The length of the telescope between the two lenses equals the sum of their focal lengths. You see all this in the sketch to the right.

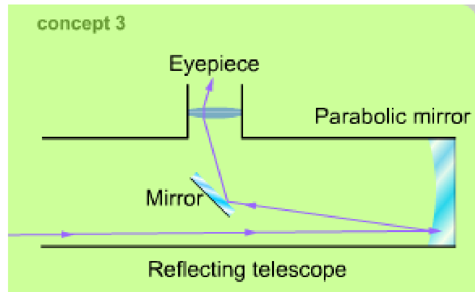
As you can see in Concept 2, the image produced by the objective lens is inverted, smaller than the object, and real. The image produced by the eyepiece from this real image subtends a larger angle than the object, and it is virtual. The larger angle is an example of angular magnification, provided by the “magnifying glass” of the eyepiece. In this case it is being used to view the real image created by the objective lens, rather than a concrete object. The eyepiece does not invert the real image created by the objective lens, so the final image is inverted compared to the original object.

The telescope configuration we have been discussing works well for astronomy. The image it produces is inverted, but this typically poses no problem for astronomers, since they are accustomed to seeing telescope-rendered views of the objects they study. However, telescopes used to study terrestrial objects (like birds) use other configurations, including perhaps a diverging lens or a set of prisms, in order to avoid presenting an inverted image to the eye.

A different type of telescope designed for astronomy is called a *Newtonian reflector*. “Newtonian” refers to the developer of this type of telescope, Sir Isaac Newton, and “reflector” refers to the use of mirrors. In a reflector, a relatively large concave mirror performs the “light collecting” and first image creation duties that the objective lens performs in a refracting telescope. You see a diagram of a reflector to the right. Light passes down the tube, is collected by a curved mirror, redirected by a flat mirror and then passes through the eyepiece.



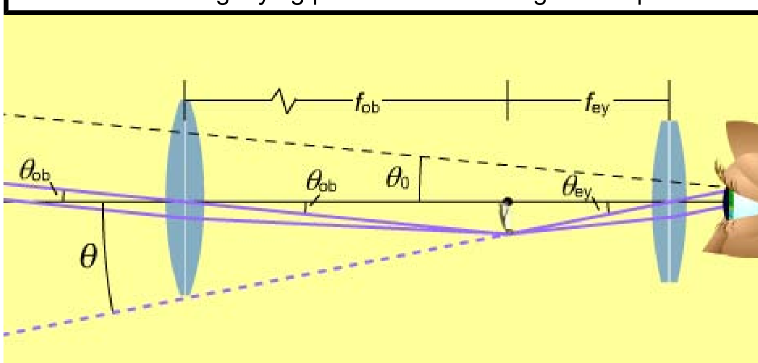
How it works
Two converging lenses
Focal points at same location
Final image inverted, at infinity, virtual



Reflecting telescope
Uses mirror in place of objective lens

Why mirrors? Consider the reflecting telescope at California’s Mount Palomar Observatory, which is famous for its 200 inch reflector. A glass objective lens of the same diameter would have a mass of about 40,000 kg (and weigh about 50 tons), more than three times the mass/weight of the mirror actually used. The weight of the mirror is still huge, but reducing the mass by more than two-thirds is a worthwhile accomplishment. In addition, lenses are subject to chromatic aberration (different colors of light focusing at different distances), while mirrors are not. A mirror is also easier to support (since it can be supported from behind), and only one side of the mirror needs to be manufactured with optical precision, versus the two sides of a lens. These factors play out in the construction of large research telescopes: The largest refractor in existence, at the Yerkes Observatory in Wisconsin, is only a little over one meter in diameter, while the largest reflectors, the Keck telescopes, are nearly ten meters in diameter.

37.23 - The magnifying power of a refracting telescope



equation 1

Angular magnifying power

$$M = -\frac{f_{ob}}{f_{ey}}$$

M = angular magnification
f_{ob} = focal length of objective lens
f_{ey} = focal length of eyepiece

Angular magnification is defined as the ratio of angular image size to angular object size. This definition applies to telescopes as well as to magnifying glasses, but developing a useful equation to calculate magnification is a distinct exercise for telescopes, based on how they are used. To fast forward to the punch line, the angular magnification of a refracting telescope is calculated by dividing the negative of the objective lens’s focal length by the eyepiece’s focal length. You see this in Equation 1.

The equation makes it clear that the more powerful eyepieces are the ones with shorter focal lengths. An example of this equation “at work” is supplied by amateur telescopes that have a single objective lens but can accommodate several interchangeable eyepieces, each one labeled with its “magnifying power.” The negative sign in the equation indicates that the final virtual image presented by the telescope to the eye is inverted.

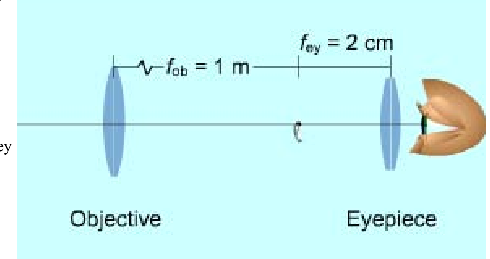
In the rest of this section, we discuss the need for and derivation of this equation. In the derivation of the magnification for a simple magnifier we determined the angle subtended by the object by placing it at the point of optimal naked-eye inspection: the near point of the eye. However, telescopes are used to view objects that are much farther away, so we cannot use the same technique here.

Since the object is far away, we assume that all rays arriving from it are parallel. This means the angular size θ_0 of the object as seen by the naked eye equals the angle θ_{ob} subtended by the object at the center of the objective lens, as shown in the diagram above.

The objective forms the first image at the common focal point of the two lenses. We show two rays that come from that real image and pass through the eyepiece into the eye. The angular size θ of the final virtual image viewed by the eye equals the angle θ_{ey} subtended by the real image at the eyepiece.

In the following steps we derive the equation for the angular magnifying power of a refracting telescope.

example 1



Find the magnifying power of this telescope.

$f_{ob} = 1$ m

$f_{ey} = 2$ cm $= 0.02$ m

$$M = -\frac{f_{ob}}{f_{ey}}$$

$$M = -\frac{1\ \text{m}}{0.02\ \text{m}}$$

$M = -50$ (image inverted)

Variables

angular naked-eye size of object	θ_0
angular size of image seen through telescope	θ
angle of ray through center of eyepiece	θ_{ey}
angle of ray through center of objective lens	θ_{ob}
height of real image produced by objective lens	h_i
focal length of objective lens	f_{ob}
focal length of eyepiece	f_{ey}
angular magnifying power of telescope	M

Strategy

1. Express the naked eye angular size of a distant object in terms of properties of the telescope lenses and their images.
2. Express the angular size of the virtual image seen by the eye in terms of properties of the telescope lenses and their images.
3. Divide the two angular sizes determined in the previous steps to state the angular magnification of the telescope. Simplify the result.

Physics principles and equations

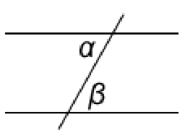
The rays from a distant object to a telescope are both parallel and paraxial.

The definition of the angular magnifying power of a telescope includes a negative sign, because the final image is inverted.

$$M = -\frac{\text{angular size of inverted image}}{\text{angular size of object}} = -\frac{\theta}{\theta_0}$$

Mathematics principles

When a transverse line cuts through two parallel lines, the "alternate interior" angles α and β illustrated in the diagram below are equal.



For the small angle θ between a paraxial ray and the principal axis of the telescope, measured in radians,

$$\theta \approx \tan\theta$$

Step-by-step derivation

We state both the naked-eye object angular size and the final virtual image angular size in terms of lengths, using the small angle tangent approximation. We divide the resulting expressions and simplify.

Step	Reason
1. $\theta_0 = \theta_{ob}$	alternate interior angles
2. $\theta_{ob} \approx \tan\theta_{ob} = -h_i/f_{ob}$	small angle approximation
3. $\theta = \theta_{ey}$	inspection
4. $\theta_{ey} \approx \tan\theta_{ey} = -h_i/f_{ey}$	small angle approximation
5. $M = -\dfrac{\theta}{\theta_0}$	definition of angular magnification
6. $M = -\dfrac{-h_i/f_{ey}}{-h_i/f_{ob}}$	substitute equations 1, 2, 3, and 4 into equation 5
7. $M = -\dfrac{f_{ob}}{f_{ey}}$	simplify

37.24 - The magnifying power of a compound microscope

Microscope eyepieces with various angular magnifying powers.

In this section, we discuss the magnifying power of microscopes. We set aside the definition used with simple magnifiers and telescopes, which was based on the ratio of the angular size of the image created by the optical instrument to the angular size of the object viewed without assistance. We do so because the naked eye angular size of an object under a microscope is near zero.

The definition used for the magnifying power of a microscope is the product of the lateral magnification associated with the objective lens and the angular magnification of the eyepiece. Microscopes often have three objective lenses on a rotating stage and a set of interchangeable eyepieces. The manufacturer labels the power of the objectives and the eyepieces. Multiplying an objective and eyepiece value yields the magnifying power of that combination.

A microscope's objective lens produces a real image whose lateral height is greater than the lateral height of the object by a certain factor. The ratio of these lateral heights defines the magnification provided by this lens. The image produced by the objective lens is further magnified by the eyepiece. Since the virtual image is essentially at infinity, the eyepiece magnification is measured in terms of an increase in angular size. This definition of magnification for a microscope is shown as the first equation on the right. To stress a point: it is the product of a magnification measured in terms of height (the objective lens factor) and magnification measured in terms of angular size (the eyepiece lens factor).

The second equation on the right provides a formula that allows you to approximate the overall magnification of a microscope based on physical characteristics of its lenses. This definition relates to the ray-tracing diagram in Equation 1, where the strength of each lens is defined by its focal point instead of its magnification. As you can see, the object is placed just outside the focal point of the objective lens, and the initial real image appears far from the objective, slightly inside the focal point of the eyepiece. The distances away from the two focal points are exaggerated in the diagram.

In the second equation, L is the distance between the two lenses and N is the near point of the human eye, with 25 cm a common value for this quantity. As the equation indicates, the magnification of the microscope increases with L, and is greater when the focal lengths of the lenses are short.

The following derivation shows that the computational formula on the right is equivalent to the definition.

Variables

	eyepiece	objective
magnification (angular, lateral)	M_{ey}	m_{ob}
focal length	f_{ey}	f_{ob}
image distance (virtual, real)	$D_i \approx \infty$	d_i large
object distance	D_o (small)	d_o (small)

distance between the lenses	L
near-point distance of observer	N

Strategy

1. Use the lateral magnification formula for the objective lens, and two rough equalities from the diagram to get an approximation for m_{ob}.
2. Combine a simple-magnifier equation with the thin lens equation to state an approximation for M_{ey}.
3. Multiply the two approximations obtained in previous steps to obtain a formula for the overall magnification.

Physics principles and equations

The equation for lateral magnification is

$$m = -\frac{d_i}{d_o}$$

The equation for the angular magnification of a simple magnifier is

$$M = \frac{N}{d_o}$$

The distance between the lenses of a microscope is $L = d_i + D_o$.

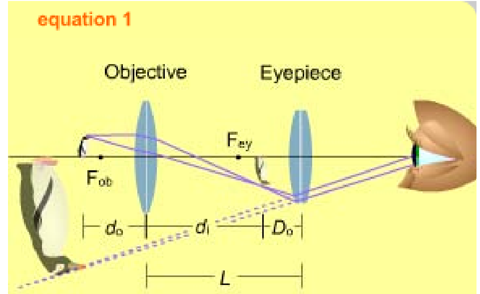
We will use the thin lens equation

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

Step-by-step derivation

First, we derive an approximation for m_{ob}.

Step	Reason
1. $L = d_i + D_o$	inspection of diagram
2. $L \approx d_i$	D_o in previous equation is small
3. $f_{ob} \approx d_o$	object close to objective focal point
4. $m_{ob} = -\dfrac{d_i}{d_o}$	lateral magnification of real image
5. $m_{ob} \approx -\dfrac{L}{f_{ob}}$	substitute approximations 2 and 3 into equation 4

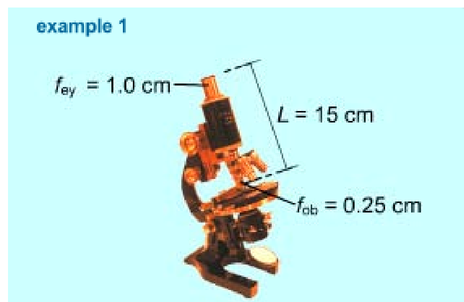


Microscope magnification

$$M = m_{ob} M_{ey}$$

$$M \approx -\frac{LN}{f_{ob} f_{ey}}$$

M = overall magnification
m_{ob} = objective lateral magnification
M_{ey} = eyepiece angular magnification
L = distance between the lenses
N = near point distance
f_{ob} = focal length of objective
f_{ey} = focal length of eyepiece



Find the magnification of this microscope. Assume *N* = 25 cm.

$$M \approx -\frac{LN}{f_{ob} f_{ey}}$$

$$M \approx -\frac{(15\text{ cm})(25\text{ cm})}{(0.25\text{ cm})(1.0\text{ cm})}$$

$$M \approx -1500$$

Now we derive an approximation for M_{ey}.

Step	Reason
6. $M_{ey} = N\left(\frac{1}{D_o}\right)$	angular magnification by eyepiece
7. $\frac{1}{D_o} = \frac{1}{f_{ey}} - \frac{1}{D_i}$	thin lens equation
8. $M_{ey} = N\left(\frac{1}{f_{ey}} - \frac{1}{D_i}\right)$	substitute equation 7 into equation 6
9. $M_{ey} \approx \frac{N}{f_{ey}}$	D_i in previous equation is large

In the final steps we multiply the two approximations obtained above and simplify the result.

Step	Reason
10. $M = m_{ob} M_{ey}$	definition of magnification
11. $M \approx \left(-\frac{L}{f_{ob}}\right)\left(\frac{N}{f_{ey}}\right)$	substitute approximations 5 and 9 into equation 10
12. $M \approx -\frac{LN}{f_{ob} f_{ey}}$	rearrange

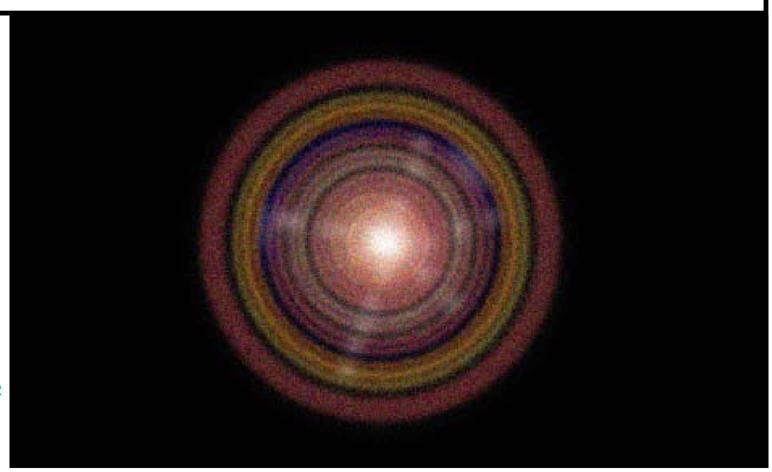
The example to the right assumes a near point of 25 cm. The microscope has a magnification of −1500. The negative sign indicates the image is inverted.

37.25 - Lens aberrations

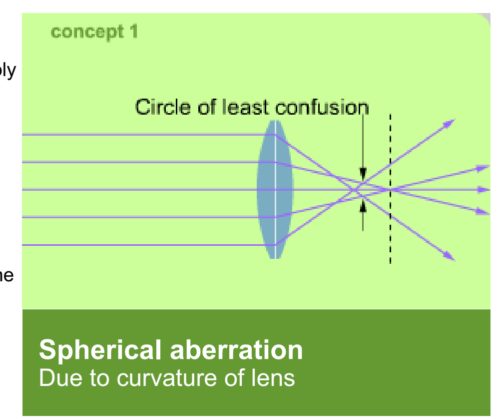
Lens aberration: Imperfection in the images formed by a lens.

Spherical aberration: Image blurring due to the spherical contour of the lens.

Chromatic aberration: Image blurring due to differing refraction of various wavelengths (colors) of light.



Spherical aberration.



You may have noticed that even when an image created by a lens is focused as sharply as possible, it is still somewhat blurry. The blurriness results from lens aberrations. Blurring means that the light rays from a single point on an object do not precisely converge at one point in the image, as the examples to the right illustrate, using horizontal rays from a point on a very distant object. A well-crafted lens causes little blur; lower quality lenses can suffer from substantial blurriness.

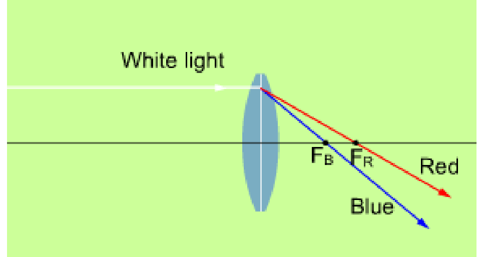
There are several types of aberration. Some aberrations may be caused by imperfections in the manufacturing of the lens, but one type arises unavoidably from the nature of a spherical lens. The lack of focus arising from spherical aberration most affects rays that strike a lens near its periphery. *Spherical aberration* means the light rays refract at an undesired angle; rather than converging to a single point, they converge to a larger region. The spherical nature of the lens causes this problem.

Instead of a single point of focus for incoming parallel rays, the lens creates a *circle of least confusion*, a region in which the rays approximately converge and an image is most satisfactorily viewed. This is diagrammed in Concept 1.

A typical image created by a spherical lens is shown above. The paraxial rays focus sharply in the center of the real image. Nonparaxial rays, having already passed through their closer focal points, are beginning to spread out again. Their contribution to the image shows up as a set of concentric rings. The outermost ring corresponds to rays passing through the outermost portions of the lens which have the shortest focal lengths.

Spherical aberration can be counteracted in a variety of ways. For instance, photographers are aware of this phenomenon, and when precise focus is important, they "stop down" the aperture of the iris that lets light into the camera lens, restricting the remaining light to the center portion of the lens. Mirrors are also subject to spherical aberration: In large astronomical telescopes, the problem is avoided by using parabolic or other non-spherical reflectors.

concept 2



Chromatic aberration
Different wavelengths refract differently

Chromatic aberration occurs because the refractive index of a material is a function of the wavelength of light. For instance, since blue light refracts more than red light when it passes from air into glass, a ray of white light disperses into its component colors when refracted by glass. This notably occurs in prisms, although there it is often the desired effect. You see a diagram illustrating chromatic aberration in Concept 2.

The photo below provides an example of chromatic aberration. There is a noticeable purple halo around the circular windows on the left side of the image. The right side of the image is actually the center of a larger cropped image. The chromatic aberration is less noticeable in this area, which corresponds to the center of the lens. Chromatic aberration worsens toward the edge of the lens.

A compound lens made with lenses of different indices of refraction can counteract this effect. This type of *achromatic lens* is found in high quality optical equipment.

37.26 - Gotchas

"Real" means the same thing with lenses as it does with mirrors. On the one hand, yes. Real images can be projected onto paper with both lenses and mirrors. On the other hand, no. With lenses, a real image is on the opposite side of the lens from the object. With mirrors, a real image is on the same side of the mirror as the object.

Signs. You need to be careful with signs when using the lens equations, and we have provided a table to help you do so.

The ray diagrams for lenses are entirely accurate and represent physical reality. The diagrams in this (and other textbooks) employ a few conventions in order to make them easier to understand: Objects are shown fairly close to lenses that are relatively thick compared to the distance to the objects. For the approximations inherent in the theory of thin lenses to hold true, the objects should be quite far from the lenses and the lenses should be much thinner. Applying these changes to diagrams would make for illustrations that stretch far across the page or computer screen.

Lenses redirect light by refraction. Converging lenses are thicker in the middle and bring light rays together. Diverging lenses are thinner in the middle and spread light out.

Lens terminology is very similar to that of mirrors. An important difference is that when we consider a single lens, virtual images appear on the same side of the lens as the object, and real images appear on the opposite side. Another difference is that lenses have a focal point on each side.

Ray diagrams for converging and diverging lenses are also similar to those for mirrors.

And like mirrors, lenses have equations that quantify the relative size, orientation and distance of the images they produce. In addition, the lensmaker's equation determines the focal length of a lens with differing radii of curvature on its two sides. Sign conventions can be trickier for lenses, so pay special attention to them.

Several lenses can be used together to enhance their magnification properties, such as in a refracting telescopes.

In a microscope, a lens called the objective creates a real image, which is used as the object for a second lens called the eyepiece. The viewer looks into the eyepiece and sees a final image that is inverted, virtual, and magnified.

The human eye contains a lens that can change shape in order to create a focused image on the light-sensitive retina at the back of the eye. A person's near point and far point are the closest and farthest distances on which she can focus.

Sometimes a lens is specified by its refractive power, which is the inverse of its focal length. Refractive power P is measured in diopters, where 1 diopter = 1 m^{-1}. This unit is commonly used with eyeglasses and contact lenses.

The angular size of an object is the angle of the field of view that is taken up by the object. It is measured in radians, which allows the use of a small-angle approximation: For small angles, the angular size of an object is approximately its height divided by its distance.

The angular magnification of a simple magnifier is the ratio of the angular size of the image to that of the object when the object is located at the near point of the human eye.

Lenses can exhibit spherical aberration just as mirrors can. They also exhibit chromatic aberration due to the differing refractive indices of different wavelengths of light. Both kinds of aberration can cause an image to look blurry.

Equations

Thin lens equation

$$\frac{1}{d_o}+\frac{1}{d_i}=\frac{1}{f}$$

Lateral magnification

$$m=\frac{h_i}{h_o}=-\frac{d_i}{d_o}$$

Lensmaker's equation, air

$$\frac{1}{f}=(n-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right)$$

Refractive power

$$P=\frac{1}{f}$$

Angular size

$$\theta\approx\frac{h}{d}$$

Angular magnification

$$M=\frac{\theta}{\theta_0}$$

Angular magnification of a magnifying glass

$$M\approx\frac{N}{d_o}=\left(\frac{1}{f}-\frac{1}{d_i}\right)N$$

$$M\approx\frac{N}{f}+1 \quad \text{(image at near point)}$$

$$M\approx\frac{N}{f} \quad \text{(image at infinity)}$$

Angular magnification of a refracting telescope

$$M=-\frac{f_{ob}}{f_{ey}}$$

Magnification of a microscope

$$M\approx-\frac{L(N)}{f_{ob}f_{ey}}$$

Conceptual Problems

C.1 An object is placed between the focal point of a converging lens and a distance twice the focal length from the lens. Draw a ray diagram to determine the nature of the image and describe the image.

i. Virtual; larger; upright
ii. Real; smaller; inverted
iii. Virtual; smaller; upright
iv. Real; larger; inverted
v. Real; larger; upright

C.2 An object is placed between a converging lens and its focal point. Draw a ray diagram to determine the nature of the image and describe the image.

i. Virtual; larger; upright
ii. Real; smaller; inverted
iii. Virtual; smaller; upright
iv. Real; larger; inverted
v. Real; larger; upright

C.3 An object is placed in front of a diverging lens. Draw a ray diagram to determine the nature of the image.

i. Virtual; larger; upright
ii. Real; smaller; inverted
iii. Virtual; smaller; upright
iv. Real; larger; inverted
v. Real; larger; upright

C.4 An object is placed farther from a converging lens than twice the lens' focal length. Draw a ray diagram to determine the nature of the image and describe it.

i. Virtual; larger; upright
ii. Real; smaller; inverted
iii. Virtual; smaller; upright
iv. Real; larger; inverted
v. Real; larger; upright

C.5 A single lens is used to create a real image. Is the image upright or inverted? Explain.

○ Upright ○ Inverted

C.6 A converging lens is placed very far away from an object, so that an image is created very near the focal point of the lens. The image is real, smaller than the object, and inverted. As you move the lens closer to the object, (a) does the image get farther away from the lens, or closer to it? (b) Does the image get larger or smaller? When the lens reaches a distance of twice the focal length from the object, (c) where is the image located? (d) What is the nature of the image at this point?

(a) ○ Farther ○ Closer
(b) ○ Smaller ○ Larger
(c)
i. Infinity
ii. At the focal point
iii. At twice the focal length
iv. At half the focal length

(d)
i. No image: rays refract parallel
ii. Real; larger than the object; inverted
iii. Real; equal size as the object; inverted
iv. Virtual; equal size; upright

C.7 If you had the urge to do so, which of the following physical quantities could you measure in hertz per diopter?

i. Frequency of light passing through a lens
ii. Time required to measure refractive power
iii. Speed of a bullet
iv. Number of wavelengths of light in focal length
v. Density of seawater

C.8 Would you expect mirrors to exhibit chromatic aberration? Why?

○ Yes ○ No

C.9 You are trying to light a leaf on fire using a magnifying glass and rays from the Sun. Where can you hold the leaf so that it lights quickest? Explain.

(a) i. Between the focal point and the lens
ii. At the focal point
iii. Beyond the focal point

(b)

C.10 Which of the following optical instruments produces a real final image (in the way it is normally used)? Explain.

- [ ] Camera
- [ ] Compound microscope
- [ ] Movie projector
- [ ] Magnifying glass

Section Problems

Section 0 - Introduction

0.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) Where must the object be relative to the focal point F to create an image on the same side of the lens as the object? (b) Where must the object be to create an image on the opposite side of the lens? (c) Where must the object be to create an upright image?

(a) i. On the opposite side of the focal point as the lens.
ii. Between the focal point and the lens.
iii. It is impossible.
iv. Anywhere.

(b) i. On the opposite side of the focal point as the lens.
ii. Between the focal point and the lens.
iii. It is impossible.
iv. Anywhere.

(c) i. On the opposite side of the focal point as the lens.
ii. Between the focal point and the lens.
iii. It is impossible.
iv. Anywhere.

Section 4 - Interactive problem: image with a diverging lens

4.1 Use the simulation in the interactive problem in this section to answer the following questions about a single object and a diverging lens. (a) Is it possible to make a real image? (b) Is it possible to make an inverted image? (c) Is it possible to make an image that is larger than the object?

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c) ○ Yes ○ No

Section 7 - Lens equations

7.1 An object is placed 23.0 cm in front of a thin lens. The image created is 5.00 cm behind the lens. (a) Is the object distance positive or negative? (b) Is the image distance positive or negative? (c) What is the focal length of the lens? (d) Is the lens converging or diverging?

(a) ○ Positive ○ Negative
(b) ○ Positive ○ Negative
(c) ________ cm
(d) ○ Converging ○ Diverging

7.2 An object is placed 23.0 cm in front of a thin lens. The image created is 5.00 cm in front of the lens. (a) Is the object distance positive or negative? (b) Is the image distance positive or negative? (c) What is the focal length of the lens? (d) Is the lens converging or diverging?

(a) ○ Positive ○ Negative
(b) ○ Positive ○ Negative
(c) ________ cm
(d) ○ Converging ○ Diverging

7.3 The focal length of a thin lens is 0.110 m. You want to use it to produce a real image at 0.600 meters from the lens. How far from the lens should the object be placed?

__________ m

7.4 The focal length of a thin lens is −9.00 cm. You place an object 15.0 cm from the lens. (a) Is the image produced real or virtual? (b) What is the image distance (include the correct sign)?

(a) ○ Real ○ Virtual
(b) __________ cm

7.5 An object is placed in front of a thin lens at a distance of 7.00 cm, and the lens produces an inverted image that is twice the size of the object. (a) Is the image height positive or negative? (b) Is the image real or virtual? (c) How far is the image from the lens? (d) What is the focal length of the lens?

(a) ○ Positive ○ Negative
(b) ○ Real ○ Virtual
(c) __________ cm
(d) __________ cm

7.6 An object is placed in front of a thin lens at a distance of 12.0 cm, and the lens produces an upright image that is one-third the size of the object. (a) Is the image height positive or negative? (b) Is the image real or virtual? (c) Is the image distance positive or negative? (d) What is the image distance? (e) What is the focal length of the lens?

(a) ○ Positive ○ Negative
(b) ○ Real ○ Virtual
(c) ○ Positive ○ Negative
(d) __________ cm
(e) __________ cm

7.7 You are making a thin lens. The near surface (the surface on the object side) has a radius of curvature of +5.00 cm and the far surface has a radius of curvature of +6.00 cm. What is the focal length of the lens if the index of refraction of the material is 1.60?

__________ cm

7.8 The thin lens equation given in this book is sometimes called the Gaussian form of the equation. Another form, called the Newtonian form, uses the distance x between the object and the first focal point and the distance x' between the second focal point and the image. Use the Gaussian form to derive the Newtonian form:

$$xx' = f^2$$

7.9 A thin lens has a near surface with a radius of curvature of −5.00 cm and a far surface with a radius of curvature of +7.00 cm. (a) Is the lens converging or diverging? (b) What is the focal length of the lens if the index of refraction of the material is 1.74?

(a) ○ Converging ○ Diverging
(b) __________ cm

7.10 A lens has two convex surfaces (both surfaces bulge at the middle). The left surface has a radius of curvature whose magnitude is 11.2 cm and the right surface has a radius of curvature whose magnitude is 13.0 cm. The index of refraction of the lens is 1.41. (a) What is the focal length of the lens for light traveling from left to right? (b) What is the focal length of the lens for light traveling from right to left?

(a) __________ cm
(b) __________ cm

7.11 The cornea and lens of the eye each refract light as it passes into the eye. We consider the cornea and lens together as one thin lens whose shape, and effective focal length, can be changed by the muscles around the eye. When incoming light is parallel, such as for an infinitely far away object, the eye muscles shape the effective lens to have a focal length of about 2.50 cm to produce a focused image on the retina. What focal length lens do the muscles cause to create a focused image of an object 35.0 cm away?

__________ m

7.12 A camera has a lens of focal length 50.0 mm and is focused on an object 0.800 m from the lens. (a) To focus on an object effectively at infinity, will the lens have to move toward or away from the film? (b) How far will the lens have to move? Give the answer in millimeters.

(a) ○ Toward ○ Away
(b) __________ mm

7.13 A small light bulb is 1.50 meters from a screen. You have a converging lens with a focal length of 0.250 m. There are two possible distances from the bulb at which you could place the lens to create a sharp image on the screen. What is the larger of these distances?

__________ m

7.14 A small light bulb is placed a distance *d* from a screen. You have a converging lens with a focal length of *f*. There are two possible distances from the bulb at which you could place the lens to create a sharp image on the screen. (a) Derive an equation for the distance *z* between the two positions that includes only *d* and *f*. (b) Use this equation to show that the distance *d* between an object and a real image formed by a converging lens must always be greater than or equal to four times the focal length *f*.

Section 9 - Interactive problem: focus a camera

9.1 Use the information given in the interactive problem in this section to calculate the correct lens-film distance so that the image will be in focus. Test your answer using the simulation.

__________ cm

Section 10 - Interactive problem: optical bench with a lens

10.1 Use the information given in the interactive problem in this section to calcuate (a) the object distance and (b) the focal length of the lens in order to create the desired image. Test your answer using the simulation.

(a) __________ cm
(b) __________ cm

Section 14 - Nearsightedness

14.1 Nellie is nearsighted. She cannot focus on objects farther than 40.0 cm from her unaided eye. (a) Does she require a converging or diverging lens to correct her vision? (b) What focal length must her corrective contact lens have to bring into focus the most distant objects?

(a) ○ Converging ○ Diverging
(b) __________ cm

Section 17 - Refractive power: diopters

17.1 What is the refractive power of a lens whose focal length is 0.648 meters?

__________ diopters

17.2 What is the refractive power of a lens whose focal length is −0.286 meters?

__________ diopters

17.3 The refractive power of a lens is −2.50 diopters. What is the focal length?

__________ m

Section 18 - Angular size

18.1 A quarter is 2.3 cm across. What is its approximate angular size in radians when viewed from 3.5 meters away?

__________ radians

18.2 A penny is 1.8 cm across. What is its approximate angular size when viewed from 1.6 meters away? State your answer in degrees.

__________ °

18.3 You measure a crater on the moon to have an angular size of 3.64×10^{-4} radians. The distance to the moon is 3.84×10^{8} m. What is the approximate diameter of the crater?

__________ m

18.4 The Hubble Space Telescope has a resolution of about 2.4×10^{-7} radians for visual wavelengths. (a) If you were to view a quarter at a distance such that it subtended the same angle, how far away would it have to be? A quarter is 2.3 cm in diameter. (b) The resolution of the Hubble at infrared wavelengths is about 9.7×10^{-7} radians. How far away would a quarter have to be to subtend this angle?

(a) __________ m
(b) __________ m

Section 19 - Angular magnifying power: simple magnifier

19.1 An entomologist examining an insect adjusts the position of her magnifying glass so that the image of the insect is at her near point, which is 31.5 cm from her eye. If the angular magnifying power of the magnifying glass is 3.75, what is the approximate distance between the insect and the magnifying glass?

__________ cm

19.2 You are using a magnifying glass with a focal length of 4.85 cm to examine a rare coin. You place the magnifying glass so that the image of the coin is at your near point of 25.0 cm. What is the approximate angular magnifying power of the magnifying glass?

____

Section 23 - The magnifying power of a refracting telescope

23.1 A large refracting telescope has an objective lens with focal length 0.85 m. (a) You use an eyepiece with focal length 0.016 m. What is the angular magnification? (b) You now swap the eyepiece for one with a longer focal length of 0.030 m. Will the magnitude of the angular magnification increase or decrease? (c) What is the new angular magnification?

(a) __________
(b) ○ Increase ○ Decrease
(c) __________

23.2 You have a converging lens with a focal length of 0.85 m, and wish to construct a refracting telescope with an angular magnification of −45. (a) Should you use the lens you have as the objective or the eyepiece? (b) To finish the telescope, what should be the focal length of the second lens?

(a) ○ Objective ○ Eyepiece
(b) __________ m

23.3 You have a refracting telescope consisting of an objective lens with focal length 0.550 m and an eyepiece with focal length 0.0160 m. You view Jupiter through this telescope when it has an angular diameter of 2.23×10^{-4} radians on the sky. (a) What is the angular magnification of your telescope? (b) What is the angular size of the image of Jupiter that you see through the telescope? Hint: You may use the definition of angular magnifying power given earlier in the chapter.

(a) __________
(b) __________ radians

Section 24 - The magnifying power of a compound microscope

24.1 A microscope has an objective lens focal length of 0.30 cm, an eyepiece focal length of 1.3 cm, and the two lenses are separated by 14 cm. Using an average nearpoint of 25 cm, what is the approximate overall magnification of the microscope?

____

Additional Problems

A.1 A lensmaker creates a lens with a material whose index of refraction is 1.43. The near surface has a radius of curvature of −12.0 cm and the far surface has a radius of curvature of +14.0 cm. What is the refractive power of the lens?

__________ diopters

A.2 A compound microscope provides an overall magnification of −155. The focal length of the eyepiece is 2.30 cm. Assume an average near point of 25.0 cm. What is the approximate lateral magnification of the objective lens? Recall that the image produced by a compound microscope is at an infinite distance.

____

A.3 Elizabeth is nearsighted. Without glasses, she can see objects clearly when they are between 15.0 cm and 90.0 cm away from her eyes. Her glasses are designed to be worn 2.00 cm from her eyes, and have a focal length so that objects at infinity produce images at her far point. When she is wearing these glasses, how **close** to her eye can an object be before it appears out of focus?

__________ cm

A.4 A man who wears eyeglasses that have a power of −3.75 diopters wants to get contact lenses. His glasses sit 1.90 cm from his eyes, but the contacts will sit directly on his eyes. What will the power of the contact lenses have to be?

__________ diopters

A.5 A simple refracting telescope that produces a final upright image can be constructed with a converging objective lens and a diverging eyepiece (a design used by Galileo). The tube length is set at the objective focal length plus the (negative) eyepiece focal length for viewing far away objects (the lenses are at the extreme ends of the tube). (a) Where is the final image produced by this telescope located? (b) Is the image real or virtual? (c) If the magnification of a particular telescope of this design is 4.00, and the length of the tube is 30.0 cm, what is the objective lens focal length?

(a) i. At one of the objective focal points
ii. At one of the eyepiece focal points
iii. At infinity

(b) ○ Real ○ Virtual

(c) ________ cm

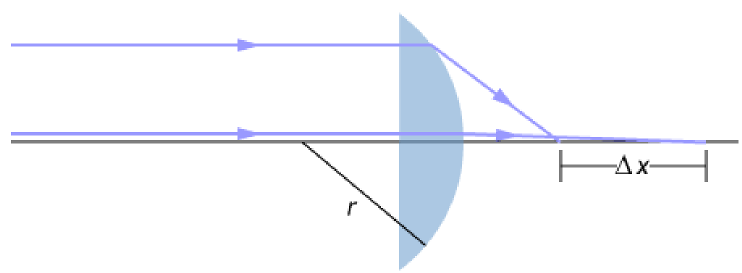
A.6 People who suffer from cataracts may have the lenses in their eyes replaced by artificial lenses. (a) If a given person has a distance of 22.0 mm between her lens and retina, what refractive power would an artificial lens need to have to allow her to see far away objects clearly? (b) There is no accommodation with the artificial lens, so she has to use glasses to focus on objects close-up. If an object is held 25.0 cm away from her glasses, what is the refractive power of the glasses she will need to see the object clearly?

(a) ________ diopters

(b) ________ diopters

A.7 (a) Derive the lensmaker's equation for a lens of index of refraction n_2 that is surrounded on one side by a material with an index of refraction of n_1 and on the other side by a material with an index of refraction of n_3. Hint: Start with the single surface equation and use the techniques in the derivation of the lens equations. (b) Check that the equation reduces to the lensmaker's equation for air when $n_1 = n_3 = 1$. (c) Derive the thin lens equation for the same situation. (d) Check that your equation reduces to the thin lens equation for air.

A.8 A particular lens is flat on one surface and spherically convex (bulges at the middle) on the other, with a radius of curvature whose magnitude is 20.0 cm. Two rays are incident parallel to the principal axis at distances of 0.400 cm and 10.5 cm, first passing through the flat surface, then refracted at the convex surface. The refracted rays experience noticable spherical aberration: They do not pass through the focal point. The index of refraction of the lens material is 1.50. Find the distance between the points where the refracted rays cross the principal axis.

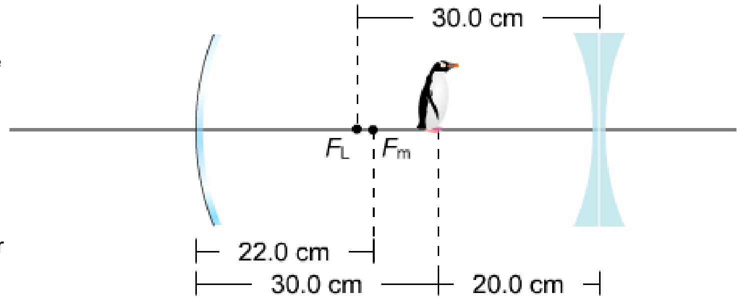


________ cm

A.9 A child requires eyeglasses with diverging lenses, but is concerned that they will look too thick. The radius of curvature of the near surface of the lenses has a magnitude of 55.0 cm and the radius of curvature of the far surface has a magnitude of 35.0 cm. If the lenses are circular with a diameter of 5.00 cm and a thickness at the center of 0.100 cm, how thick are the glasses at the edge?

________ cm

A.10 A diverging lens is placed 50.0 cm to the right of a concave mirror. An object is placed between the two, 30.0 cm from the mirror. The lens has a focal length of −30.0 cm and the mirror has a radius of curvature of 44.0 cm. Use only the light that leaves the object and hits the mirror first to answer the following questions. (a) Is the final image to the left or right of the mirror? (b) Is the final image real or virtual? (c) Is the final image upright or inverted compared to the original object? (d) How far is the final image from the mirror? Give the absolute value of this distance if it is negative. (e) What is the overall magnification of the system?



(a) ○ To the left ○ To the right
(b) ○ Real ○ Virtual
(c) ○ Upright ○ Inverted
(d) __________ cm
(e) __________

A.11 A concave spherical mirror is pointed at the Sun. Its radius of curvature is 2.50 m. The Sun subtends an angle of 0.533° in the sky. (a) What is the diameter of the image of the Sun created by the mirror? Hint: Use the magnification equations and the small angle approximation. (b) You want to use the mirror as a "solar cooker." If the intensity of sunlight incident on the mirror is 1.00×10^3 W/m^2 and you want to supply energy to a large cooking pot at a rate of 255 W, what must be the circular "cross-sectional" area of the cooker? Note that this is not the area of the curved mirror, but of the planar circle that intercepts the sunlight, the imaginary lid that caps the mirrored bowl. (c) What intensity is produced at the image?

(a) __________ cm
(b) __________ m^2
(c) __________ W/m^2

chapter 38 Interference

38.0 - Introduction

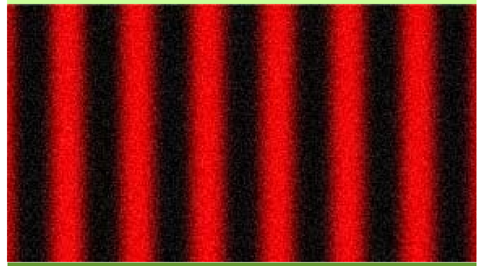
Light is a particle. Many of the great scientists of the 17th and 18th centuries who made fundamental contributions to the study of optics, including Isaac Newton, thought that light consisted of a stream of "corpuscles," or particles. In the 20th century, Albert Einstein explained the photoelectric effect. His explanation, for which he was awarded the 1921 Nobel Prize, depended on the fact that light acts like a particle. This property of light led to the coining of the term "photon" for a single particle of light by the chemist Gilbert Lewis.

Light is a wave. Between the 18th and 20th centuries, physicists discovered many wave-like properties of light. They found that a number of phenomena they routinely observed with water waves they could also observe with light.

For instance, the English scientist Thomas Young (1773-1829) showed that light could produce the same kinds of interference patterns that water waves produce. At the right, you see examples of interference patterns formed by light and by water waves. The similarities are striking. In this chapter, you will apply to light some of what you have studied about the interference of sound waves and traveling waves in strings.

Let there be light. Is light a particle, a wave, or both? Perhaps an Early Authority had it right. Light is light. It is a combination of electric and magnetic fields. Trying to classify light as a particle or as a wave may be a fruitless effort − better to revel in its unique properties. In this chapter, we will revel in its wave-like properties, and discuss the topic of interference. Your prior study of electromagnetic radiation modeled as a wave phenomenon will prove useful.

concept 1



Interference of light waves
Pattern of bright and dark on screen

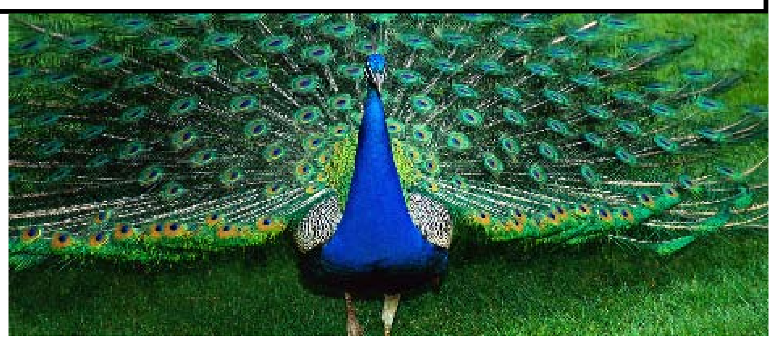
concept 2

Interference of water waves
Expanding circular ripples
Pattern of disturbance and calm

38.1 - Interference

In Concept 1, you see an *interference pattern* created by causing a beam of light to pass through two parallel slits to illuminate a viewing screen. Constructive interference of light waves accounts for the bright regions (called bright *fringes*) while destructive interference causes the dark fringes.

In this section, we review some of the fundamentals of interference, and discuss the conditions necessary for light to make the pattern you see to the right. You may have already studied the interference of mechanical waves; for instance, what occurs when two waves on a string interact. In this chapter, you will study what happens when electromagnetic waves meet. Some of the same principles and terminology are used in discussing both kinds of interference.



Grass is green because it selectively absorbs nongreen colors. In contrast, the shimmering "color" of a peacock's feathers is due to interference effects.

When two light waves meet, the result can be constructive or destructive interference. In the following discussion, we assume that the waves have equal amplitude. Constructive interference creates a wave of greater amplitude and more intensity than either source wave; destructive interference results in a wave of smaller amplitude and less intensity than either source wave. At any point in a two-slit interference pattern such as that to the right, light waves from the two sources meet and interfere constructively, destructively, or partially (exhibiting a degree of interference somewhere between complete constructive and destructive interference).

concept 1

Interference pattern
Bright and dark fringes

To create an interference pattern, a physicist needs light that is:

1. *Monochromatic*. This means light with a specific wavelength. For instance, experimenters can produce the pattern you see in Concept 1 by using pure red light.
2. *Coherent*. This means the phase difference between the light waves arriving at

any location remains constant over time.

The first condition, monochromatic light, means that all the light passing through the slits must have the same wavelength. White light, which is a mixture of many wavelengths (colors), does not produce distinct interference patterns. We often used a similar condition when analyzing the interference of mechanical waves, restricting our attention to what occurred when two waves of the same wavelength met.

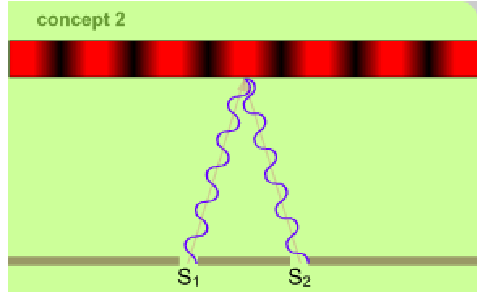
We also implicitly used the second of the two conditions stated above with waves on a string. The phase difference between the waves sometimes remained constant throughout the string, and over long time periods.

For light waves, coherence can be achieved by causing light from a single point source to pass through two narrow slits, separating an initial beam of light into two beams with a constant phase difference. You see what happens when coherent light emerges from two slits and falls on a surface in Concepts 2 and 3.

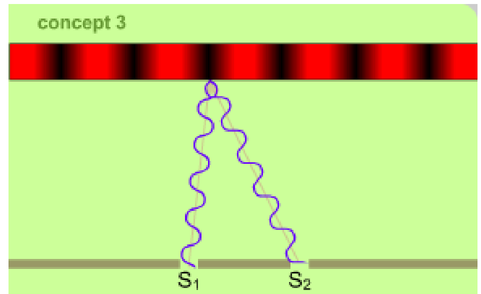
If the light were not coherent, we would not be able to see interference. If the phase differences varied over time, then there would be constructive, partial, and destructive interference occurring at every particular spot at different times, and no overall pattern of interference could be readily observed. For example, if the coherent light from the two slits were replaced with incoherent light emanating from two light bulbs, only a uniform glow of illumination would be visible on the viewing screen.

For complete constructive interference to occur, two light waves have to be in phase, meaning there is zero phase difference (or else the phase difference is an integer multiple of 2π radians, or $360°$). At the point shown in Concept 2 the waves have no phase difference, peak meets peak and trough meets trough, and there is complete constructive interference. This causes the maxima, or regions of maximum intensity in the interference pattern.

In Concept 3, you see the result at a different location on the viewing screen where light waves from the two slits meet out of phase, with a phase difference of π radians, or $180°$ (the phase difference could also be an odd integer multiple of π radians, such as 3π radians, -5π radians and so on). With this phase difference, the result is complete destructive interference. Peak meets trough, trough meets peak, and the waves cancel. Complete destructive interference causes the minima, the stripes of darkness in the interference pattern.



Constructive interference
Waves in phase
Result: bright region (fringe)



Destructive interference
Waves out of phase
Result: dark region (fringe)

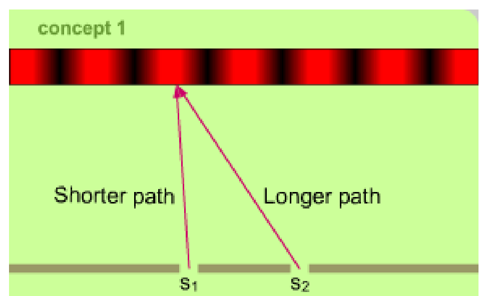
The regions in between the lightest and darkest points are the result of partial interference. Here, the waves are somewhat out of phase, so they do not reinforce each other completely, nor do they completely cancel. The visual result is a progression of shades of intensity between the bright fringes caused by complete constructive interference and the dark fringes caused by complete destructive interference.

38.2 - Double-slit experiment: wavelike nature of light

The English scientist Thomas Young explored interference and diffraction in a series of experiments that demonstrated the wavelike nature of light. Young performed his first, crucial experiment in 1801. Here, we provide a summary of what he did. His actual procedure was slightly more complex than the description below due to the pioneering nature of his equipment.

Young shined light at a pair of barriers containing slits to create the interference pattern you see on the right. The first barrier had a single slit that acted as a coherent point source of light. (This single-slit barrier is not shown in the diagrams.) The coherent light then traveled through two parallel slits in the second barrier, each the same distance from the single slit, and then reached a viewing screen. In the diagrams you see the double-slit barrier and the pattern of interference he observed on the screen.

The pattern was one of equally spaced bright and dark fringes. Young knew that water waves passing through a pair of slits could cause a similar interference pattern. The fact that both water waves and light produced exactly the same type of pattern supported his hypothesis that light acted as a wave. Young further reinforced his position when he demonstrated that he could use interference patterns to calculate the wavelength of the light that he used in the experiment.



Young's double-slit experiment
Interference pattern appears on screen
Shows wavelike nature of light
Path lengths of rays to screen differ

Why do waves create the interference pattern you see here? To answer this question, let's consider the diagram in Concept 2. Two rays of light meet at a point that is the same distance from each slit. The drawings of the rays emphasize their wavelike nature.

In this case, the two rays intersect the center of the screen in phase because both travel the **same distance** to this point. They were in phase when they passed through the slits, and since they traveled the same distance, they remain in phase at their point of convergence. They constructively interfere and produce a bright area of illumination. Other bright fringes occur to either side of the central fringe where the path difference between the waves is not zero, but one full wavelength, two wavelengths, and so on.

Now look at Concept 3. The rays of light reaching the viewing screen intersect at the first dark fringe to the left of center. They travel **different distances** to the surface: The ray from S_2 travels a half wavelength farther than the one from S_1. Because they start in phase at their respective slits, this means the waves are $\pi/2$ radians $(180°)$ out of phase when they arrive at the screen. This causes destructive interference, and darkness.

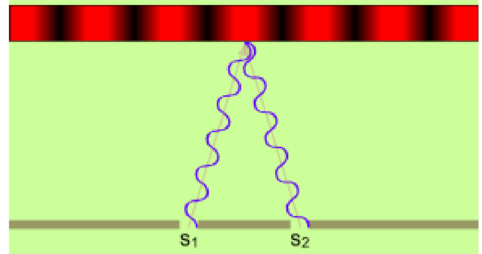
The pattern of bright and dark fringes extends to both the left and the right on the screen. The light is interfering constructively at the bright fringes, and destructively at the dark fringes, because of different path lengths to these regions and the resulting phase differences.

There are a few limitations to showing Young's apparatus in a compact diagram. First, the diagram is far from being drawn to scale. The screen should be much farther from the double-slit barrier than we show here, and the slits should be narrower and closer together. In actual interference experiments, the interfering rays from the two slits are practically parallel. Second, we vastly exaggerate the wavelength of the light.

You may have a question about what you would see if you conducted this experiment yourself. What if, at some instant, two waves meet at the screen and are in phase, but their electric and magnetic fields both happen to be zero at that point? Would you see "flickering" as the two reinforcing waves moved from peak to trough and back again? The answer is no: The frequency of light is so great that you only perceive the average brightness of a region; the human eye does not perceive changes in intensity due to the oscillation of a light wave.

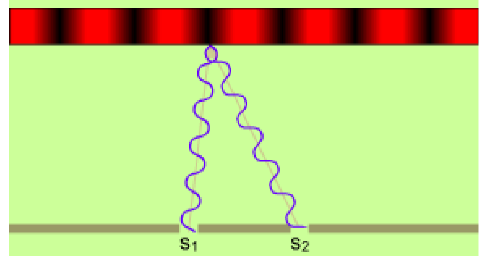
You do not even perceive flicker in systems oscillating at far lower frequencies, much less than the frequency of visible light, which is on the order of 10^{14} Hz. For example, a computer monitor refreshes its display 60 times a second, but you do not ordinarily perceive any flicker when you look at it.

concept 2



Pattern of bright and dark
Constructive interference: bright fringes

concept 3



Pattern of bright and dark
Destructive interference: dark fringes

38.3 - Double-slit experiment: wavelength of light

In his double-slit experiment, Young was able to demonstrate that light acts like a wave by exhibiting the interference pattern produced by coherent monochromatic light passing through parallel slits. He did more than that, however. He was able to use his experimental apparatus to calculate the wavelength of the light, employing an analysis that we will now discuss in some detail.

To explain his thought process, we start with a normal line drawn from the midpoint between the slits S_1 and S_2 to the viewing screen, as shown in Concept 1. Each fringe in the interference pattern lies at an angular displacement θ from this normal line. The two rays that interfere to form the fringe have a path length difference that Young was able to relate to the angle, the wavelength of the light, and the distance between the slits.

The bright fringe in the middle of the screen, equidistant from the two light-transmitting slits, is called the *central maximum*. You see this fringe labeled $m = 0$ in the diagrams to the right.

There are dark fringes on either side of the central maximum, also labeled $m = 0$, caused by destructive interference. Bright fringes labeled $m = 1$ exist on either side of the first pair of dark fringes.

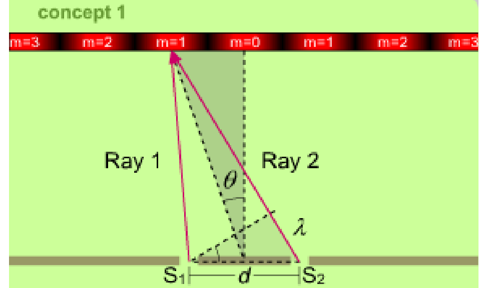
What is the relationship between the wavelength of the light and the bright fringes? The two light rays travel the same distance to the central maximum. Their paths to this point have no difference in length; their phase difference is 0 rad.

The centers of the next pair of bright fringes, the bright fringes to the immediate left and right from the central maximum, occur when the path length difference of the rays equals one wavelength, λ. Their phase difference is 2π rad. To visualize why these rays interfere constructively, picture two sinusoidal waves that are in the same place, having zero phase difference. If one of them is shifted along its path by exactly one wavelength, its peaks and troughs will still coincide with those of the other wave.

This discussion establishes an important premise: There is a relationship between the locations of the bright fringes and wavelength. Next, we will use trigonometry to show how the wavelength of monochromatic light can be determined from quantities that Young could measure empirically.

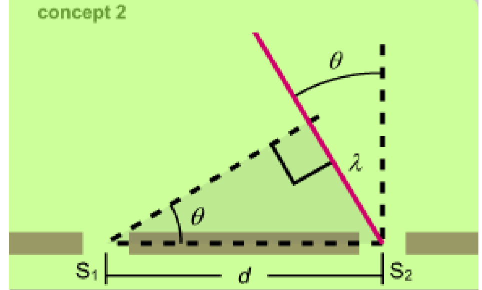
The analysis depends on an approximation. In Concept 1 (and Equation 1), we show two rays converging at $m = 1$. We want to use a single angle to describe the direction of these rays even though they are not perfectly parallel. We use the angle θ, whose left

concept 1



First bright fringe
Path difference is one wavelength λ

concept 2



Relating path length difference to wavelength
Angle to fringe is θ
In the right triangle
· Base angle is θ
· Opposite side is path difference λ
· Hypotenuse is slit separation d
· Trigonometric ratio is $\sin\theta = \lambda/d$

side is the line from the midpoint between the slits to the fringe $m = 1$.

The diagram is not to scale, which makes the approximation look worse than it actually is. The double-slit barrier and viewing screen are shown hundreds of times closer together than they would be in a typical experiment. If the distance between the two were faithfully depicted, it would show how close both rays are to being parallel to the left side of the angle θ.

In Concept 2 we zoom in on the slits and construct a right triangle that relates θ, the wavelength λ of the light, and the distance d between the slits. We do so by drawing a dashed line from slit S_1 perpendicular to the ray that is emerging from slit S_2. We have already described how the angle between this ray and the normal to the viewing screen is very close to θ, and we have labeled it this way in the diagram. Since the angle between any two lines equals the angle between their perpendiculars, θ is also a base angle in the right triangle.

The length of the side of the triangle opposite θ is very nearly equal to the extra length traveled by ray 2 as it propagates from slit S_2 to the center of the first bright fringe $m = 1$, as you can verify by glancing back at Concept 1. In a real experiment, this approximation is extremely accurate, and we are justified in labeling the opposite side of the triangle as λ, the wavelength of the light. (Recall that for the bright fringe $m = 1$, the path length difference is exactly λ.)

Since the hypotenuse of the right triangle depicted in Concept 2 is the distance d between the slits, trigonometry tells us that $\sin\theta = \lambda/d$. We state this relationship in Concept 2.

Given its extremely small size, Young could not directly measure the wavelength of light, but he could observe and measure the angles to the bright fringes and the distance between the slits in his experiments. In other words, he could use θ and d to determine λ indirectly. In effect, he created a "magnifying glass" with which to examine the wavelength of light. Young's experiment provided the first experimental way to determine the wavelength of light.

Returning to the present day: As mentioned before, the central bright fringe and the dark fringes immediately to either side are numbered as $m = 0$. The fringes beyond them (both bright and dark) are referred to as $m = 1$, $m = 2$, and so on, as the diagrams to the right reflect. The non-central bright fringes occur when the path difference is a multiple of the wavelength of the light. Shifting a wave by two full wavelengths, or three, or four, has the same interference effect as shifting it by one wavelength. Using this fact, and constructing right triangles just like the one above, except that their opposite sides are 2λ, 3λ, 4λ and so on, we arrive at the general bright fringe formula shown in Equation 1.

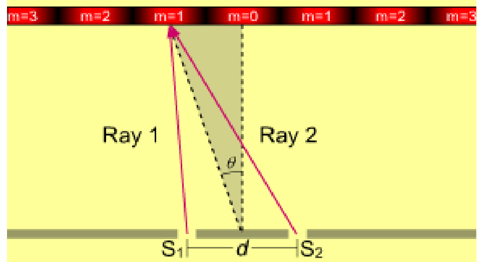
We apply a similar analysis to dark fringes. In the illustration for Equation 2, we show the paths of the light rays that interfere to form one of the "zeroeth" dark fringes, a first minimum of intensity. Again, keep in mind that in a properly proportioned diagram rays 1 and 2 would be practically parallel, with their direction to the screen accurately measured by the angle θ. We also show the equation for dark fringes.

To explain this equation, we consider the relationship between destructive interference, phase difference and path length. Complete destructive interference occurs at the first minimum when the waves are $\pi/2$ out of phase, which occurs when the path length difference is $\lambda/2$. Why? Imagine taking two waves that are in phase, peak meeting peak, and shift one by $\lambda/2$. Peak will now meet trough, and the result is complete destructive interference, or darkness.

The other dark fringes are explained in an analogous fashion. The second pair of dark fringes, labeled $m = 1$ (and lying between the $m = 1$ and $m = 2$ bright fringes), occurs when the path difference is $1\frac{1}{2}\lambda$, the third occurs at $2\frac{1}{2}\lambda$, and so on. We can construct right triangles similar to the one we used for the bright fringes to arrive at the formula in Equation 2.

The example problem describes an experiment with monochromatic light of a yellowish green hue. Using the angle θ and the slit separation, an experimenter was able to calculate the wavelength of the light. If the viewing screen were 10 meters from the slits, the first bright fringe would be about 3 millimeters from the central maximum.

equation 1



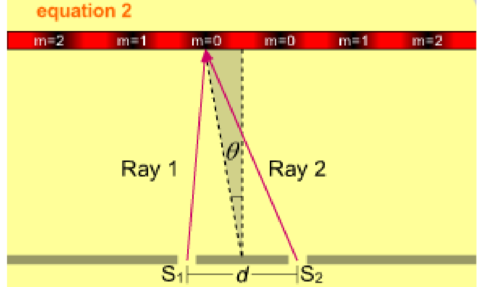
2-slit interference: BRIGHT fringes

*m*th bright fringe:

$$\sin\theta = \frac{m\lambda}{d}; \qquad m = 0, 1, 2, \ldots$$

θ = angle from normal to fringe
m = fringe number, λ = wavelength
d = distance between slits' midpoints

equation 2

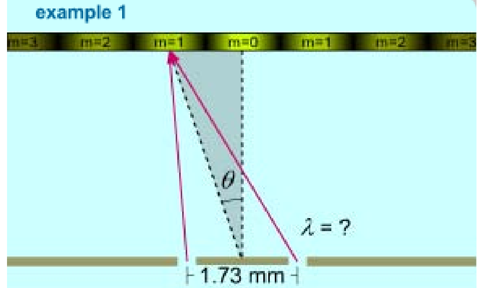


2-slit interference: DARK fringes

*m*th dark fringe:

$$\sin\theta = \frac{(m+\frac{1}{2})\lambda}{d}; \qquad m = 0, 1, 2, \ldots$$

example 1



What is the wavelength of this yellowish-green light? The angle *θ* is 0.0190°.

$\sin\theta = m\lambda / d$

$\lambda = (d)(\sin\theta) / m$

$\lambda = (1.73\times10^{-3}\text{ m})(\sin 0.0190°) / 1$

$\lambda = 5.74\times10^{-7}\text{ m} = 574\text{ nm}$

38.4 - Double-slit experiment: white light

Interference patterns are often depicted with black-and-white diagrams or photographs. However, it is important to keep in mind that such experiments are actually performed with monochromatic light, not white light. A collection of distinct bright and dark fringes cannot really be obtained with white light. In this section we explain why.

You see interference patterns created by red, "white" and blue light in Concept 1. By "white," we mean the everyday light that you perceive as white light, made up of light of many different wavelengths. When you see these wavelengths individually, they appear as the various colors of the rainbow. When they are all combined together as *polychromatic light*, your brain interprets the light as white.

In Concept 1, we show the interference pattern for monochromatic red light on top, with the bright fringe $m = 0$ in the center. On the bottom, we do the same for monochromatic blue light. (We picked red and blue because they are at opposite ends of the visible light spectrum.) In between these two patterns we show an interference pattern, produced by white light, that is the "sum" of the patterns of the individual colors.

Where there are substantial regions of both red and blue, as in the central bright fringes, the colors "sum" to white. If you pick a region in the combined (middle) white pattern that is reddish, you will note that there is a red bright fringe directly above in the red pattern, but the blue interference pattern below has a dark fringe at the same position. A dark fringe means there is no blue light at the location, and since there is red light, the region is reddish. A similar argument applies for bluish regions in the white-light pattern.

As with monochromatic double-slit interference patterns, the white light pattern has a central maximum. Light of every wavelength interferes constructively there, since it travels the same distance from either slit to reach this point. Then, as you can see in the first few bright fringes on either side of the central maximum, the differing interference patterns of the monochromatic fringes add to make a rainbow pattern. Beyond these first few, the contributing fringes due to the various component wavelengths of white light are so far offset from each other that the combined pattern is hopelessly blurred. (Although we have described portions of the white light interference pattern as rainbowlike, their cause, interference, is quite different from the cause of rainbows in the sky: refraction and reflection.)

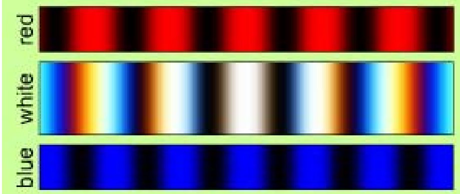
These effects can be analyzed mathematically as well. To do this, we use the equation shown in Equation 1, relating the wavelength λ of monochromatic light to θ, the angle of the m^{th} bright fringe in its interference pattern. This is the same as an equation stated in a prior section, slightly rearranged to emphasize the role of the wavelength of the light in determining $\sin\theta$.

Looking at the equation, you can see that the greater the wavelength, the greater the angle between the normal and each bright fringe of the pattern. Since red light has a longer wavelength than blue, you can predict that its pattern will be more "spread out" than the pattern for blue. In fact, since $\sin\theta \approx \theta$ for the small angles typically encountered in interference experiments, for a particular experimental configuration, the "stretching" of the pattern is just about proportional to the wavelength.

The illustration in Equation 1 is a conceptual diagram showing a side-by-side comparison of interference patterns for red and blue light. We show the results of two experiments. In Experiment 1, blue light was shined through the slits, and we only show the results of that experiment on the left side. In Experiment 2, the same experiment was conducted with red light, and we only show the results of that experiment on the right side. Since the wavelength of the blue light in the experiments is 72% of the wavelength of the red light, the angle θ to the first bright blue fringe is about 72% of the size of the angle θ' to the first bright red fringe.

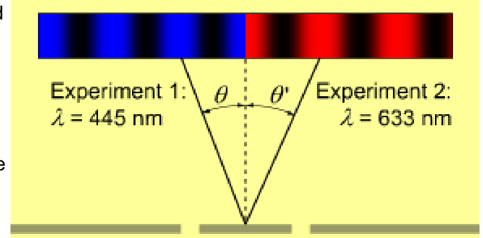
Remember that this is a conceptual diagram, with the viewing screen shown close to the double slits so you can see both the screen and the slits. The angles θ and θ' are quite small for actual equipment. In a typical experiment the wavelength of the light ranges from 400 to 700 nm, while the much larger distance between the slits is a fraction of a millimeter (equal to hundreds of thousands of nanometers). This means that λ/d, and with it $\sin\theta$, is a small value, which means the angle itself is very small.

concept 1



Double slit experiment: white light
"White" light composed of all colors
Overlapping patterns give rainbow effect

equation 1

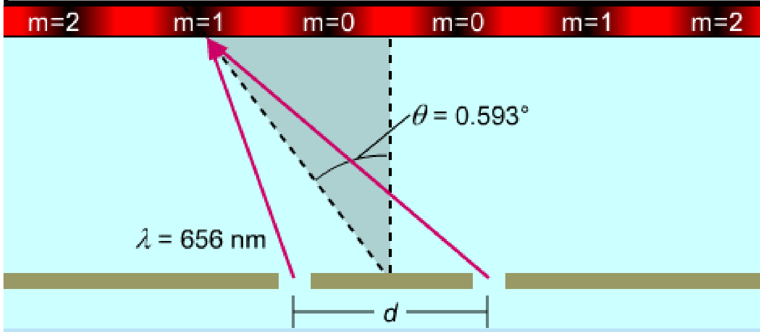


Angle to *m*th bright fringe

$$\sin\theta = \left(\frac{m}{d}\right)\lambda$$

θ = angle from normal to fringe
m = 0, 1, 2,…
d = distance between slits
λ = wavelength of light

38.5 - Interactive checkpoint: double-slit experiment



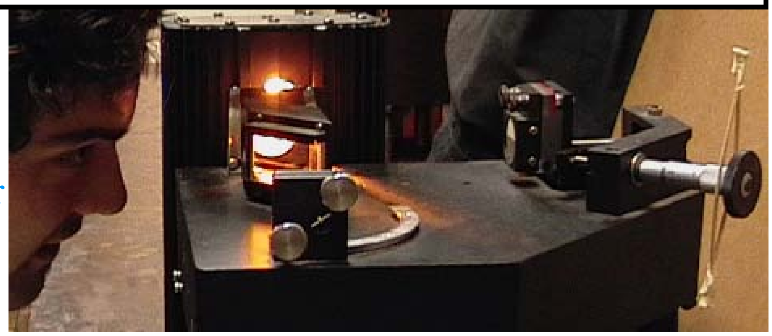
Monochromatic light of wavelength 656 nm is used in a double-slit experiment. The (exaggerated) angle from the centerline to the dark fringe indicated on the viewing screen is measured to be 0.593°. How far apart are the slits' midpoints?

Answer:

$d =$ ______ m

38.6 - Michelson interferometer

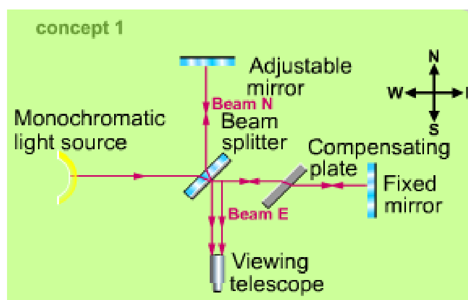
Interferometer: A device that uses the interference of two beams of light to make precise measurements of their path difference. Historically used to study the nature of light.



Michelson interferometer. In this view the telescope is on the left, the adjustable mirror is on the right, and the fixed mirror is in front.

Albert A. Michelson, an American physicist, is famous for conducting an experiment that helped Albert Einstein to develop his theory of special relativity. Michelson's experiment helped to disprove the existence of the "ether," an invisible medium that many scientists believed was required for the transmission of light waves. In this section, we discuss the design of one piece of equipment he used, now called the Michelson interferometer, and how it uses the interference patterns of known wavelengths of light to measure minute differences in the path lengths of two beams of light.

You see a photograph of an interferometer above, and a conceptual diagram to the right. We use compass directions like "north" to explain the directions of the various beams. In the center of the diagram is a *beam splitter*. It is a plate of glass whose back is coated with a thin layer of silver. The beam splitter reflects half the light that falls on it, and lets the rest pass through.



Michelson interferometer
Used for precise length measurements
Interference patterns change with path difference

The interferometer works in the following fashion:

1. Monochromatic, coherent light strikes the beam splitter from the west side.
2. The initial beam splits into two at the beam splitter. The silver coating reflects beam N to the north, and the splitter refracts beam E, which passes to the east.
3. Beam N travels to an adjustable mirror, reflects back, and then passes back through the splitter to a viewing telescope.
4. Beam E goes through a compensating plate of glass, reflects off a fixed mirror, passes through the compensating plate again and then reflects directly off the back of the splitter to reach the viewing telescope.

This system causes two mutually coherent beams of light to travel by different paths. An operator can control the length of the north-south path by changing the location of the adjustable mirror. The compensating plate ensures that both beams pass through the same amount of glass, which is important since light slows down in glass.

The purpose of this equipment is to control and detect with extreme precision changes in the path length difference between the two beams of light. Michelson knew the relationship between path length differences and interference patterns. A path difference of half a wavelength would cause the beams to interfere destructively, and a path length difference of an entire wavelength would cause complete constructive interference.

To measure a tiny increment of path length, Michelson could place a bright fringe at the center of the image created in the viewing telescope by setting the distance of the adjustable mirror. When he then moved the adjusting mirror a very small distance using a finely threaded screw adjustment, the image in the telescope would shift so that a dark fringe would be at the center. He knew he had moved the mirror one-quarter wavelength because the light traveled that additional distance to the adjusting mirror and then traveled back, making for one-half wavelength difference in total. A half wavelength path difference is the difference between complete constructive and complete destructive interference. In actual practice, Michelson's experiment had a few more "tricks" to it than we have described here, but this section conceptually summarizes how it worked.

38.7 - Phase changes in reflection

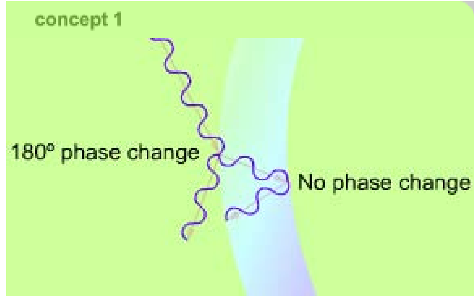
At the right, you see a ray of light traveling through air and encountering a soap bubble. Some of the light reflects off the front surface of the bubble substance, and some passes through and reflects when it reaches the back surface. The light changes phase $180°$ when it reflects off the first surface, and it does not change phase when it reflects off the second surface.

In general, a wave changes phase by π radians $(180°)$ when it reflects off a surface with a higher refractive index than the medium it is traveling in. It does not change phase when it reflects off a material with a lower refractive index than the medium it is traveling in. This accounts in part for the interference phenomenon that makes soap bubbles exhibit a rainbow of colors.

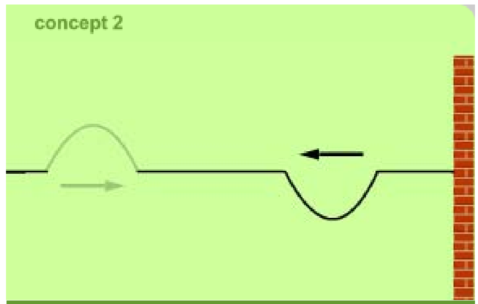
Why does light reflect in this way? We will explain, using the analogy of a reflected wave on a string. In Concepts 2 and 3 you see two strings, one connected to a fixed point, and the other to a ring that is free to slide up and down. When the wave reflects off the fixed point, it changes phase by $180°$. A peak becomes a trough, and vice versa. On the other hand, when the wave reflects off the sliding ring, the reflected wave has the same phase as the incident wave.

These two results − a phase change of $180°$, and no phase change − are also observed in light waves reflected off the surfaces between two materials. In the case of light, the differences occur as a result of the relationship of the indices of refraction of the materials, and the corresponding speed of light in those materials.

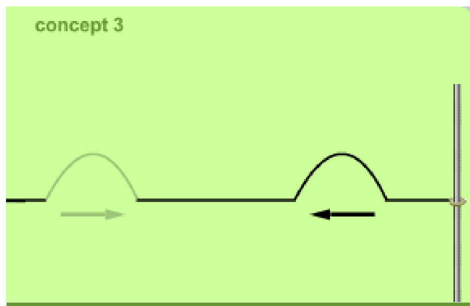
Why these differing phase changes occur with the reflection of electromagnetic waves is a digression we will not undertake here. It is a topic you will encounter if you proceed in your study of physics. Here, we note that it is empirically true: Experiments show that light reflects off a surface with a higher refractive index with a phase change of π radians. Light reflects off a surface with lower refractive index with no phase change.



Reflection and refractive indices
Reflecting off material with higher *n*
· Phase changes 180°
Reflecting off material with lower *n*
· No phase change



Wave reflecting at fixed point
Phase changes 180°



Wave reflecting at sliding ring
No phase change

38.8 - Thin-film interference

Thin-film interference: The interference caused by light waves reflecting off the two different surfaces of a thin film.

In soap bubbles and in thin layers of gasoline or oil floating on water, you sometimes see "rainbows," light of all the colors of the spectrum. The causes of rainbows in the sky and rainbows in soap bubbles are

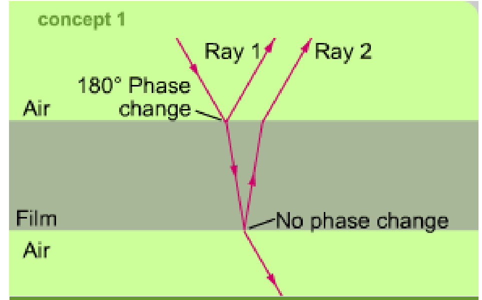
Thin-film interference in soap bubbles.

quite different. Those celestial "Pot-of-gold" rainbows are caused by refraction and reflection. Soap-bubble rainbows are caused by a phenomenon called thin-film interference.

To understand thin-film interference, we start with the fact that when light reflects off a material with a higher index of refraction than the medium it is traveling in, it changes phase by $180°$. When it reflects off a material with a lower index of refraction than the medium it is traveling in, there is no phase change.

Specifically, when light reaches the front surface of a thin film surrounded by air, some of it reflects, and changes phase by $180°$. Some of the light passes through this first surface, and then reflects off the far surface of the film. No phase change occurs here. As the light moves between the two media, it refracts as well.

This combination of reflection and refraction is illustrated in Concept 1. You see a downward ray striking the film from the upper left. Some of it reflects when it reaches the film's top surface, and we call that reflected ray 1. Some of the light passes into the film, reflects at the bottom surface, and passes back through the film again. We call that ray 2.



Thin-film interference
Interference "rainbow" caused by:
· change of phase on first reflection
· difference in travel distance of rays
· variations in thickness of film

The initial ray of light becomes two rays that have a somewhat complicated relationship in terms of their phase difference. Ray 1 changes phase $180°$ as it reflects, and ray 2 does not change phase as it reflects but it travels an extra distance through the film that ray 1 does not. Depending on the path length difference, which equals roughly twice the thickness of the film, the two rays could end up completely in phase, completely out of phase, or somewhere in between.

Interference depends not just on path difference but on wavelength as well. The type of interference occurring at a specific point in a thin film will differ by the local thickness of the film and the wavelength, or color, of the light. White light has components consisting of many colors, and in a film of variable thickness these components will interfere differently at different points.

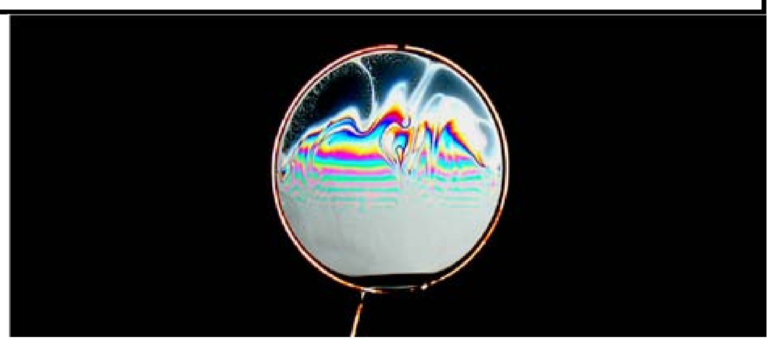
What about the rainbow of colors you can observe on a soap bubble or oil sheen? The film of a soap bubble can be thicker or thinner at various locations. The path length difference at a certain point on a soap bubble may cause destructive interference for red light, but constructive interference for blue light. This means you perceive this region as blue. At another point, the bubble will have a different thickness, and the path length difference may cause the opposite result, and you see red there. These patterns can change quickly. A slight breeze or the flow of soap to the bottom of the bubble will cause parts of it to change thickness, resulting in a new pattern.

A region of a bubble that is about to burst is often thin enough that all wavelengths of light destructively interfere there, and the bubble appears dark. The path length difference through this part of the film is essentially zero, so the change of phase accompanying reflection at the front surface, which affects all wavelengths equally, causes destructive interference.

38.9 - Thin-film equations

Thin-film interference occurs because of a phase shift caused by reflection and the different paths traveled by light. In this section, we quantify the effect, using the diagrams to the right that show a close up view of a thin film of thickness t. We assume that the light rays are nearly perpendicular to the surface so that we can approximate the extra distance traveled by ray 2 as twice the thickness of the film.

At point A, where rays 1 and 2 separate, ray 1 experiences a $180°$ phase change. At point B, ray 2 reflects, but experiences no phase change. At point C, ray 2 has traveled an extra distance essentially equal to $2t$, twice the thickness of the film.



**This soap film varies in thickness and produces a rainbow of colors.
The top part is so thin it looks black. All colors destructively interfere there.**

We will assume that points A and C are the same distance from an observer viewing rays 1 and 2, or close enough that their placement does not introduce an additional path difference between them and the observer. This is a reasonable assumption for thin films viewed from a typical distance.

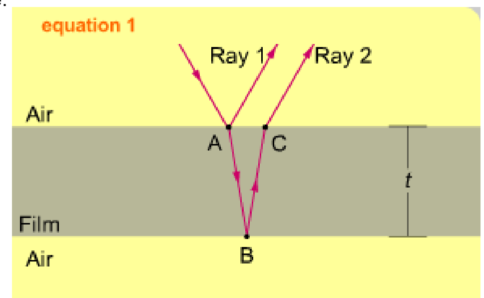
The two different types of reflection introduce a phase difference of π radians. If ray 2 travels a whole number of wavelengths during its trip A-B-C, an additional phase difference of 2π, 4π, or so on is introduced, which does not affect the type of interference. The two rays will remain a half cycle out of phase, and destructively interfere.

In equation form, this means that the condition for destructive interference is $2t = m\lambda_{\text{film}}$. The thickness of the film is t, and m is a whole number. We use the wavelength λ_{film} of the light inside the film, which is shorter than the wavelength in air. The additional equation $\lambda_{\text{film}} = \lambda/n$ (discussed in the study of refraction), where n is the index of refraction of the film and λ is the wavelength of the light in a vacuum, allows us to state the thin-film equation for destructive interference in a more convenient form: $2nt = m\lambda$. This is shown in Equation 1. (The case $m = 0$ of this equation corresponds to a film so thin $(t \approx 0)$ that all wavelengths destructively interfere, as you can observe in the photograph above.)

The condition for constructive interference is found in a similar way. For constructive interference to occur, rays 1 and 2 must be in phase when they leave A and C, so ray 2 must travel an extra distance of $(m + ½)\lambda_{\text{film}}$ within the film. Again using the equation relating wavelengths and the film's index of refraction, we obtain the equation for constructive interference shown in Equation 2 on the right.

The equations presented here apply to a thin film in air. However, they also apply whenever two media with lower indices of refraction surround a thin film with a higher index. For example, they could be applied to a layer of gasoline $(n = 1.40)$ floating on water $(n = 1.33)$ because both the air above it and the water below it have a lower index of refraction than the gasoline. Ray 1 would still undergo a phase change and ray 2 would not. This is what you see in the photograph to the right, which shows a film of gasoline on a wet roadway.

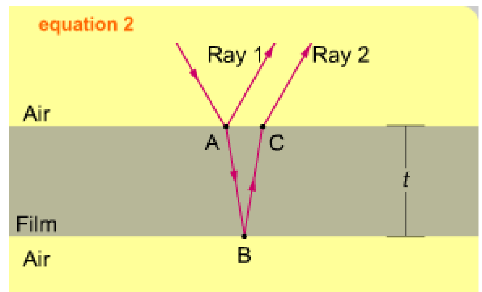
However, if the gasoline were on glass, which has a higher index of refraction (n = 1.50), ray 2 would also undergo a phase change upon reflection, so the conditions for constructive and destructive interference would be reversed and you would apply Equation 2 to analyze the minima of light reflected from the film, and Equation 1 to analyze the maxima.



Thin film: destructive interference

$$2nt = m\lambda$$

n = refractive index of film
t = thickness of film
m = 1, 2, 3, …
λ = light wavelength in vacuum (air)

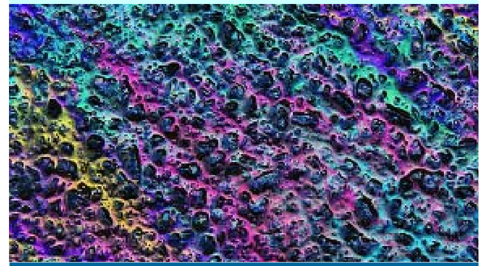


Thin film: constructive interference

$$2nt = (m + ½)\lambda$$

m = 0, 1, 2, …

example 1



A layer of gasoline (n = 1.40) on water is illuminated with white light. Find the minimum positive thickness of the gasoline layer that cannot appear red (λ = 661 nm).

$2nt = m\lambda$

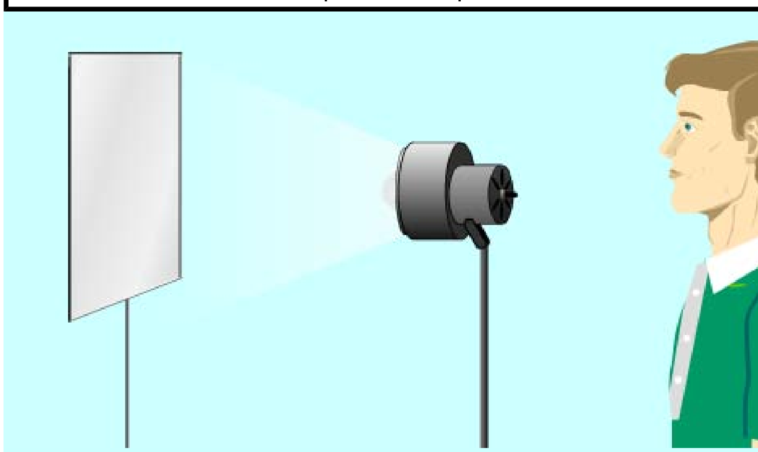
$m = 1$

$t = \lambda/2n$

$t = (661\text{ nm}) / 2(1.40)$

$t = 236\text{ nm}$

38.10 - Interactive checkpoint: a soap film in air

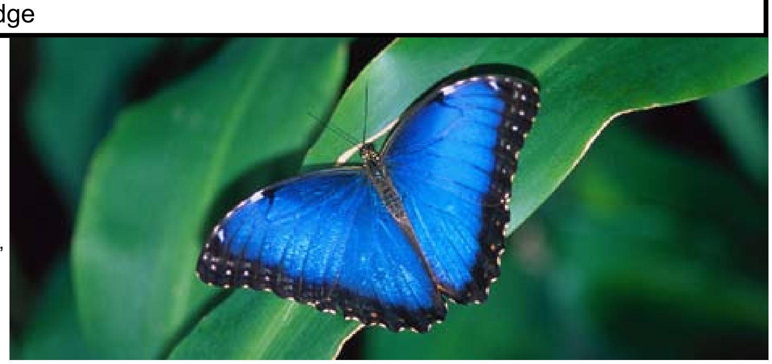


A soap film with a uniform thickness of 474 nm is surrounded by air. The index of refraction of the film is 1.33. A white light shines on the film, normal to its surface. If you are viewing the film from behind the light, what color (wavelength of light) does the film appear to be? Hint: The film will strongly reflect several distinct wavelengths of light due to constructive interference, but you will only be able to see the wavelength that is in the visible range of 400 nm to 700 nm.

Answer:

λ = ______ nm

38.11 - Thin-film interference: air wedge



The spectacular color of this *Morpho* butterfly's wings is produced by tiny scales, each enclosing an air gap between two transparent plates.

Because thin-film interference is of some interest and importance, we will work through another application of the thin-film equations.

The diagram in Concept 1 depicts an example of *air-wedge interference*. In air wedge interference, air functions as the "thin film." Two other materials with a larger index of refraction rest above and below the air, and create the wedge shape.

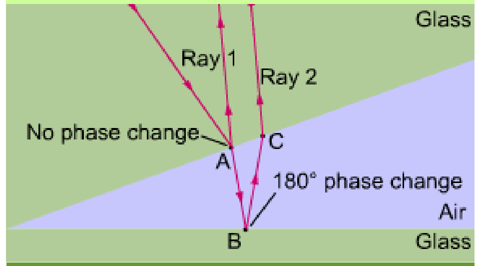
An experimenter can create this phenomenon by placing a plate of glass very close to another plate, at an angle, as shown in Concept 1. The glass surfaces should be "flat," that is, smooth on the scale of the wavelength of light. Some of the light reflects off the glass/air interface at the top of the air wedge while other light refracts through this boundary and reflects off the air/glass interface at the bottom of the wedge.

The case of the air wedge resembles the case of a thin film suspended in air, except for the location of the reflective phase change. In this case, ray 1 experiences no phase change when it reflects at A. Ray 2 changes phase by $180°$ at B and also travels farther than ray 1, going twice through the air of the wedge. This extra distance is equal to twice the thickness *t* of the wedge between B and A/C. (A and C are actually very close to each other, but the distance between them is exaggerated in the illustration for the sake of clarity.) This is the same as a soap bubble in air, except that with a bubble, the phase change occurs at the near side, not the far side, of the bubble surface.

Destructive interference occurs when ray 2 travels an integer number of wavelengths in a distance $2t$ so that it is completely out of phase with ray 1 when it reaches C. For destructive interference to occur, the thickness of the wedge satisfies the equation $2n_{air}t = m\lambda$. We can simplify this and say that $2t = m\lambda$ because the index of refraction for air is very close to 1.00. The equation for constructive interference in an air wedge is $2t = (m + ½)\lambda$.

An air wedge will create an interference pattern. Since the wedge's thickness varies, the distance the light travels while in the wedge varies depending on where it strikes it. Variations in the corresponding path length difference mean that for any particular color of light, destructive and constructive interference occur at different places along the length of the wedge. The smaller the angle between the sheets of glass that form the wedge, the more "spread out" the color bands of the interference pattern will appear. If you press two glass microscope slides between your fingers, as you see in the illustration of Equation 1, you can vary the pressure on the slides and watch the interference pattern spread and shrink.

concept 1



Thin-film interference: air wedge
Material with smaller *n* forms wedge
Phase shift occurs at second boundary
Interference pattern along length of wedge

equation 1

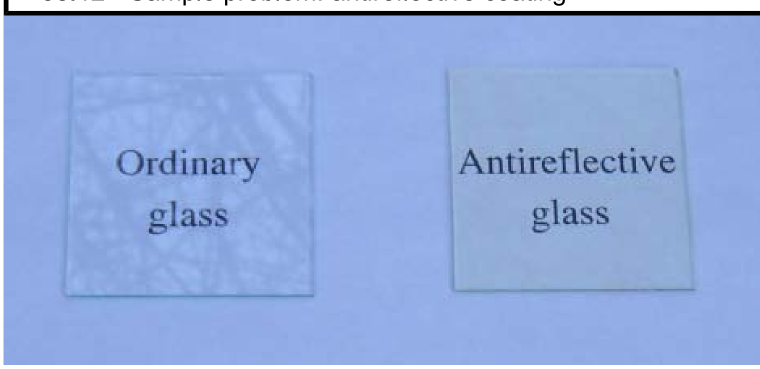
Thin-film interference: air wedge

$2t = m\lambda$ (destructive)
$m = 1, 2, 3, \ldots$

$2t = (m + ½)\lambda$ (constructive)
$m = 0, 1, 2, \ldots$

t = thickness of the air wedge
λ = wavelength in vacuum (air)

38.12 - Sample problem: antireflective coating



The antireflective glass shown is covered with a thin layer of magnesium fluoride (*n* = 1.38). The index of refraction of the glass is 1.52. What minimum thickness of magnesium fluoride is needed to cause destructive interference of light whose wavelength is in the middle of the visible range at 555 nm?

You may have seen (or not seen!) antireflective coatings which reduce reflected glare from the picture frame glass that protects artwork and from camera or eyeglass lenses. Magnesium fluoride (MgF_2) is one such type of coating. In the photograph above, the ordinary glass reflects a tangle of tree branches overhead, giving a less clear view of the caption underneath it than does the antireflective glass on the right.

This application of thin films is different from those you studied in previous sections. In this case, the film is adjacent on one side to a material with a higher index of refraction (the glass) and on the other side to a material with a lower index of refraction (air).

Variables

index of refraction of MgF_2 coating	$n = 1.38$
wavelength of light	$\lambda = 555$ nm
minimum thickness of coating	t
value of m for minimum thickness	$m = 0$

What is the strategy?

1. Consider what equation can be applied to this situation. It is not the same as the thin film in air and air wedge cases: Here, the reflected rays at both interfaces undergo $180°$ phase changes.
2. Because two phase changes occur, reverse the equations for constructive and destructive interference in a thin film in air.

Physics principles and equations

The equations used in previous sections for interference with thin films and air wedges assumed that a reflective phase change occurs at just **one** of the boundaries between materials. In this case however, both a ray reflected at the air/coating boundary and a ray reflected at the coating/glass boundary undergo phase changes. This is because each ray encounters a material with a higher index of refraction than the one in which it is traveling.

Destructive interference occurs when

$$2nt = (m + ½)\lambda$$

Step-by-step solution

Step	Reason
1. $2nt = (m + ½)\lambda$	destructive interference, two phase changes
2. $t = \dfrac{(m + ½)\lambda}{2n}$	solve for t
3. $t = \dfrac{(0 + ½)(555\text{ nm})}{2(1.38)}$ $t = 101$ nm	evaluate

38.13 - Gotchas

In a string, complete destructive interference occurs when the peaks of one wave meet the troughs of another wave with the same amplitude. The same is true with light. This is correct. With a string, the result is zero displacement. With light, the result is darkness.

A double-slit interference pattern is caused by differences in the path lengths traveled by the light emanating from each slit. This is true.

If the path length difference in a double-slit interference pattern equals one wavelength, the result is complete destructive interference. No, the result is complete constructive interference. When the path difference is one wavelength, the waves are in phase: Peak meets peak, and trough meets trough. A half wavelength path difference will cause destructive interference.

38.14 - Summary

Light exhibits the properties of both a particle and a wave. The interference patterns created by light can be explained by treating it as a wave.

Interference patterns consist of alternating bright and dark bands called fringes. To create an interference pattern, you need at least two sources providing light that is both monochromatic (having only one wavelength) and coherent (light from the different sources has a phase relationship that does not change over time). You also need a screen on which to view the pattern.

One way to create two coherent sources of light is to pass light from an ordinary source through a mask containing a narrow slit. Since the slit acts like a point source, the light that passes through is coherent.

The bright fringes in a two-slit interference pattern are the result of completely constructive interference of the light from the two sources, while the dark fringes are created by completely destructive interference.

Whether a given point on a viewing screen will be a point of constructive, destructive, or intermediate interference depends on the difference in the path lengths from each of two slits to that point.

Constructive interference results when the difference in path lengths is a whole-number multiple of the wavelength of the light creating the pattern. Destructive interference results from a path length difference equal to a half-integer multiple of the wavelength, such as $\lambda/2$, $3\lambda/2$, or $5\lambda/2$.

In a typical interference apparatus, the spacing of the fringes in the interference pattern is roughly proportional to the wavelength of the light. For this reason, white light does not produce a distinct interference pattern, since what appears on the viewing screen is the sum of the differently spread-out patterns from each of the wavelengths making up the white light.

Although we speak in terms of bright and dark fringes, an interference pattern does not consist of discrete bands of light. The intensity of the light actually varies sinusoidally along the viewing screen.

Equations

Two-slit interference: bright fringes

$$\sin\theta = \frac{m\lambda}{d}, \qquad m = 0, 1, 2, \ldots$$

Two-slit interference: dark fringes

$$\sin\theta = \frac{(m + ½)\lambda}{d}, \qquad m = 0, 1, 2, \ldots$$

Thin-film interference

Destructive: $2nt = m\lambda$

$m = 1, 2, 3, \ldots$

Constructive: $2nt = (m + ½)\lambda$

$m = 0, 1, 2, \ldots$

Air-wedge interference

Destructive: $2t = m\lambda$

$m = 1, 2, 3, \ldots$

Constructive: $2t = (m + ½)\lambda$

$m = 0, 1, 2, \ldots$

Phasors are a useful tool for analyzing interference patterns, especially those caused by multiple-slit barriers.

An interferometer is an instrument that takes advantage of interference to make precise measurements of length. It relies on the fact that a specific difference in the path lengths of two monochromatic coherent beams of light causes a specific interference pattern at a point where the beams meet, and a microscopic change in the path difference causes an easily visible change in the interference pattern.

When a light wave reflects from a material with a higher index of refraction than the one in which it is traveling, it experiences a $180°$ phase change. When a light wave reflects from a material with a lower index of refraction, there is no phase change. This affects the nature of thin-film interference, where waves that reflect from the front surface of a thin film interfere with other waves that refract through the front surface and then reflect from the back surface.

Chapter Assumptions

Unless stated otherwise, use the following indices of refraction

n_{air} = 1.00

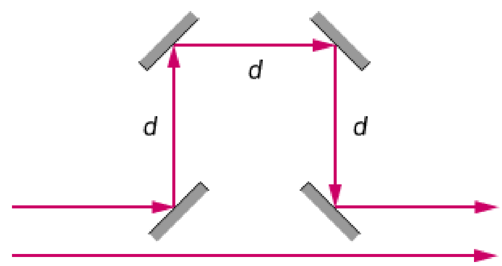
n_{water} = 1.33

n_{glass} = 1.50

Conceptual Problems

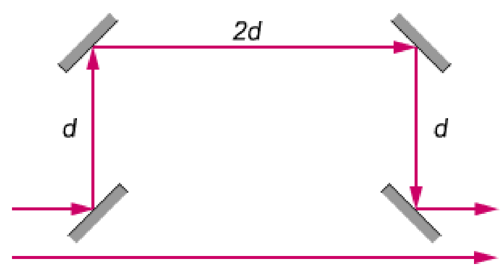
C.1 Two beams of light, initially in phase, follow the paths shown. In terms of the wavelength λ, what is the minimum length *d* that will result in complete constructive interference between the two beams?

○ $\lambda/6$ ○ $\lambda/3$ ○ $\lambda/2$ ○ $3\lambda/4$



C.2 Two beams of light, initially in phase, follow the paths shown. In terms of the wavelength λ, what is the minimum length *d* that will result in complete destructive interference between the two beams?

○ 4λ ○ $\lambda/8$ ○ $\lambda/4$ ○ $3\lambda/2$



C.3 In the old days before cable television, everyone received broadcast television signals by using an antenna perched on the television or on the roof. You might think that if you had a clear line of sight to a nearby broadcasting station, you would always receive a strong signal, but this is not true. Explain why the presence of surrounding buildings, which do not block the signal but do reflect some of it, might reduce the signal received at a television antenna with a clear line of sight.

C.4 In recent years, cancer researchers have proposed a new way of treating previously inoperable brain tumors. Some malignant tumors can be destroyed by focusing a narrow, high-energy x-ray beam on them. The beam must have a high power, but at high intensities, a single beam will destroy healthy tissue in the path to the tumor. At lower powers, a single beam will keep the healthy cells alive, but it will not kill the cancer cells. In order for the treatment to be effective, a tumor must receive a full concentration of energy at one instant in time, but healthy cells must never receive the full concentration. How can this be done?

C.5 A beam of coherent light is split into two beams, beam 1 and beam 2. Beam 1 travels a distance D_1 in arriving at point P and beam 2 travels a shorter distance D_2 in arriving at the same point. If these two beams undergo complete destructive interference at P, the path difference $D_1 - D_2$ must equal which of the following.

i. An odd number of half-wavelengths
ii. Zero
iii. A whole number of wavelengths
iv. A whole number of half-wavelengths

C.6 A pair of double slits is cut into a thin aluminum barrier, and coherent laser light passes through the slits. The interference pattern is observed on a faraway screen. Some ice is placed in contact with the bottom of the aluminum barrier so that it slowly cools. Thermal contraction causes the aluminum plate to become shorter in all of its linear dimensions. What happens to the interference fringes?

i. The fringes stay absolutely fixed.
ii. The outlying fringes move closer to the centerline.
iii. The outlying fringes move farther from the centerline.
iv. The fringes disappear.

C.7 A double-slit interference experiment using red light is set up in a vacuum chamber, from which all of the air has been pumped out. As air is slowly let back into the chamber, what happens to the interference fringes?

i. The fringes stay absolutely fixed.
ii. The outlying fringes move closer to the centerline.
iii. The outlying fringes move farther from the centerline

C.8 State whether the fringe spacing (measured on the screen) in a two-slit interference pattern increases, decreases, or stays the same if each of the following changes is made. Assume that each change is made independent of the others and that the apparatus is returned to its original state before the next change is made. (a) The slit separation is decreased. (b) The entire experiment is immersed in water. (c) The screen is placed farther away from the two slits. (d) The color of the monochromatic light used in the experiment is changed from green to orange.

(a) i. Fringe spacing increases
ii. Fringe spacing decreases
iii. Fringe spacing is unchanged

(b) i. Fringe spacing increases
ii. Fringe spacing decreases
iii. Fringe spacing is unchanged

(c) i. Fringe spacing increases
ii. Fringe spacing decreases
iii. Fringe spacing is unchanged

(d) i. Fringe spacing increases
ii. Fringe spacing decreases
iii. Fringe spacing is unchanged

C.9 If you look inside a container of gasoline, it appears as colorless as water does. However, once you place a droplet of it on, say, a water surface and it spreads out so that it becomes very thin, the gasoline begins to reflect incident light in bright rainbow patterns. The equation for constructive interference has no physical limitations requiring the thickness *t* of the film to be very small. Why must the film be in fact very thin before colors start to appear in the reflected light?

C.10 You are the manager of an optical engineering department. One of your engineers develops a very thin coating, designed to eliminate reflected light for one color. It is designed for use on a type of glass whose index of refraction is slightly higher than that of the thin coating. Excitedly, she claims that a film made of the same material and thickness may be applied to any surface whatsoever, and work in the same way. Is her claim correct? Explain.

○ Yes ○ No

C.11 A droplet of oil is deposited on a still lake. Viewing it directly from above, it looks primarily green. As the oil spreads out and the layer becomes thinner, does the color initially shift toward the shorter wavelength, blue end of the spectrum, or the longer wavelength, red end of the visible spectrum? Assume that the refractive index of water is greater than that of the oil. Explain your answer.

○ Blue ○ Red

C.12 An antireflective coating is placed on a plate of glass, and white light is allowed to shine onto it. There is less energy in the reflected beam than when the coating is not present. According to the principle of conservation of energy, the difference in energy cannot simply disappear. Where does it go?

Section Problems

Section 3 - Double-slit experiment: wavelength of light

3.1 A screen is separated from a two-slit source by 1.71 m. The centers of the slits are 0.0143 cm apart. A second-order bright fringe is found 1.34 cm from the center line (that is, the center of the fringe corresponding to $m = 0$). (a) What is the wavelength of the visible light? (b) What is the distance, measured along the screen, between this bright fringe and the bright fringe corresponding to $m = 3$?

(a) __________ m
(b) __________ m

3.2 In a double-slit experiment, slits separated by 1.22e−4 m are illuminated by a coherent 571 nm light source. What is the angular position of the first bright fringe, measured from the centerline?

__________ °

3.3 A collimated beam of light whose wavelength is 495 nm falls on a screen containing a pair of long, narrow slits whose centers are separated by 0.102 mm. Determine the spacing between the two third-order bright fringes on a screen that is located 2.39 m from the apertures.

__________ m

3.4 Light of wavelength 675 nm falls on a pair of narrow slits and the third-order bright fringe is seen at an angle of 23.0 degrees from the center of the interference pattern. What is the separation between the centers of the double slits?

__________ m

3.5 The first-order bright fringe appears 0.350 cm from the centerline when a light is passed through a double-slit apparatus. The distance between the centers of the slits is 0.450 mm and the screen is 2.71 m from the pair of slits. Find the wavelength of the light.

__________ m

3.6 The hot new performance art band, Monotone, is in town. Their set consists of a sinusoidal, pure note played from two in-phase speakers that are located 4.10 m apart, symmetrically located on either side of stage center. The back wall of the concert arena is located 114 m from the speakers. At the temperature of the arena, sound travels at 343 m/s. During the first song, titled "955 Hz", latecomer Jim is going to his seat against the back wall. While he is shuffling his way along, starting from the center aisle, he suddenly reaches a spot where he cannot hear the song any more. (a) What is the wavelength of the sound? (b) How far from the center aisle is he at that moment? (And if you think this concert is far-fetched, check out the piece "Four minutes thirty-three seconds (of silence)" by the composer John Cage.)

(a) __________ m
(b) __________ m

3.7 In a double-slit experiment, slits separated by 0.223 mm are illuminated by a coherent 5.30×10^{14} Hz light source. What is the angular position of the first bright fringe, measured from the centerline?

__________ °

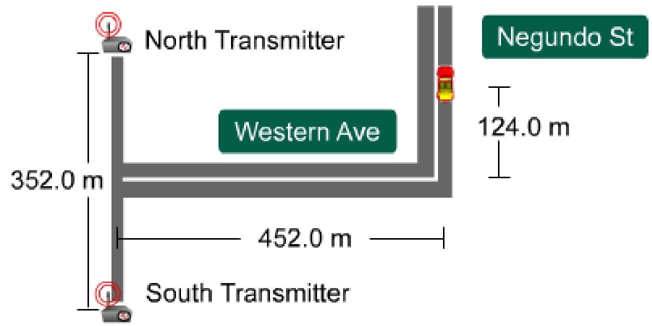
3.8 Light of frequency 4.65×10^{14} Hz falls on a pair of narrow slits, and the second-order bright fringe is seen at an angle of 17.4°. What is the separation between the double slits?

__________ m

3.9 The second-order bright fringe appears 7.35e−3 m from the centerline when a light is passed through a double-slit apparatus. The distance between the centers of the slits is 0.450 mm and the screen is 2.61 m from the pair of slits. Find the frequency of the light.

__________ Hz

3.10 Two shortwave radio antennas broadcast identical, in-phase signals at the same frequency. The transmitters are 176.0 m north, and 176.0 m south of Western Ave, respectively, as shown (that is, they are separated by 352.0 m). Western Ave is 452.0 m long. Starting at the end of that avenue, a car drives north along Negundo Street, which lies parallel to the line joining the two radio antennas. The car first encounters a minimum in reception after it travels 124.0 m.



What is the wavelength of the radio waves? Assume that the car and the transmitters are all at the same altitude.

__________ m

Section 4 - Double-slit experiment: white light

4.1 In a two-slit experiment, the slit separation is 0.301 mm, and the slits are located 1.73 m from a screen. Two distinct coherent light sources shine through the slits, one of wavelength 461 nm and another of wavelength 586 nm. The interference patterns are superimposed on the screen. What is the separation on the screen between the third-order (m = 3) bright fringes of the two patterns? Consider the third-order fringes that are on the same side of the centerline.

__________ m

4.2 A two-component beam of light, consisting only of two wavelengths 650 nm and 520 nm, is used to obtain interference fringes in a double-slit experiment. The separation between the centers of the slits is 2.43×10^{-4} m and the distance of the plane of the slits from the screen is 720 cm. What is the smallest angular displacement from the central maximum where the centers of the bright fringes due to the two colors will coincide?

__________ °

4.3 In a double-slit experiment, two parallel slits are illuminated first by light of wavelength 460 nm, and then by light of unknown wavelength. The third-order (m = 3) dark fringe resulting from the known wavelength of light falls in the same place on the screen as the second-order (m = 2) bright fringe from the unknown wavelength. What is the unknown wavelength?

__________ nm

Section 6 - Michelson interferometer

6.1 You are using a Michelson interferometer to measure the wavelength of the orange-red emission line in the spectrum of the krypton-86 atom. When the movable mirror is shifted by a distance of 303 nm, the interference between the two beams moves from constructive to destructive back to constructive. To three significant digits, what is the wavelength of this emission line?

__________ m

6.2 The adjustable mirror in a Michelson interferometer is moved a distance of 0.223 mm and 769 bright fringes cross the field of view of the telescope. What is the wavelength of light that is being used?

__________ m

6.3 A standard helium-neon laser has a wavelength of 633 nm, and is used in a Michelson interferometer experiment. If the adjustable mirror is moved a distance of 0.0216 mm, how many bright fringes cross the field of view of the telescope?

__________ fringes

6.4 The textbook speaks of moving the adjustable mirror in a Michelson interferometer by as little as a quarter wavelength of light, using a "finely threaded screw adjustment." This is an interesting claim: Just how finely threaded does such an adjustment screw have to be? Assume that the screw can be turned by as little as one-quarter of a degree with fair accuracy. How many screw threads must there be in each millimeter of the screw's length in order for a turn of one-quarter degree to advance the screw by 125 nm, a quarter of a wavelength of blue light? State your answer to the nearest hundredth. Hint: The *pitch* of any screw is defined as the distance between adjacent threads, or equivalently, the distance the screw advances after one full turn, 360°.

__________ threads/mm

Section 9 - Thin-film equations

9.1 Which of the following formulas would you use to find the thickness of a film producing complete constructive interference for reflected light? (a) Light reflects off a soap film floating in air. (b) Light reflects off a soap film on a glass plate (the glass has a higher index of refraction than the soap film).

(a) ○ $2nt = (m + 1/2)\lambda$ ○ $2nt = m\lambda$

(b) ○ $2nt = (m + 1/2)\lambda$ ○ $2nt = m\lambda$

9.2 White light illuminates an oil film on water. Viewing it directly from above, it looks red. Assume that the reflected red light has a wavelength of 615 nm in air, and that the oil has a thickness of 2.43e−7 m. What is the refractive index of the oil? Assume that the refractive index of water is greater than that of the oil.

____

9.3 What is the minimum thickness of a soap bubble needed in order for the reflected light from the outer and inner surfaces to constructively interfere? The light incident on the film has a wavelength of 655 nm. Assume that the index of refraction for the soap film is 1.35.

__________ m

9.4 A thin film of oil floats on a calm pond of water. The oil has a uniform thickness of 3.90e−7 m, and an index of refraction of 1.28. White light shines on the oil film, and a single color of visible light is seen directly above in the reflection. What is the wavelength of this color? (The wavelength of visible light is in the range 400-700 nm.)

__________ m

9.5 A thin coating is applied to a transparent plate. When viewed directly from above in white light, a single color of reflected light is visible. The coating has thickness 1.23×10^{-7} m and index of refraction of 1.35. (a) Is the index of refraction of the transparent plate greater than or less than the index of the coating? (b) What is the wavelength of the reflected light? (Visible light has wavelengths in the range 400-700 nm.)

(a) i. Greater
ii. Less

(b) __________ m

Section 11 - Thin-film interference: air wedge

11.1 A pair of very flat glass plates, 7.41 cm long, touch at one end and are separated at the other end by a small piece of 44 gauge copper wire, 5.08×10^{-5} m in diameter. An air wedge is formed between the glass plates by this supporting wire. Light of wavelength 631 nm illuminates the apparatus from above. How many bright fringes will be seen from above, along the 7.41 cm distance?

__________ bright fringes

11.2 A pair of very flat glass plates, 17 cm long, touch at one end and are separated at the other end by a piece of paper, 6.3×10^{-5} m thick. An air wedge is formed between the glass plates by this support. Light of wavelength 430 nm shines perpendicularly from above. What is the distance between the observed bright fringes?

__________ m

11.3 A narrow wedge of air is created between a pair of very flat glass plates, which touch at one end. They are separated at the other end by a small grain of sand. The angle between the plates is 4.5×10^{-4} rad, and light of wavelength 6.0×10^{-7} m illuminates the apparatus from above. How many dark interference fringes per cm are observed?

__________ fringes/cm

Section 12 - Sample problem: antireflective coating

12.1 A thin film on a lens is 1.10e−7 m thick and is illuminated with white light. The index of refraction of the film is 1.40. For what wavelength of visible light will the lens be nonreflecting? Assume the material that the lens is made of has an index of refraction greater than 1.40.

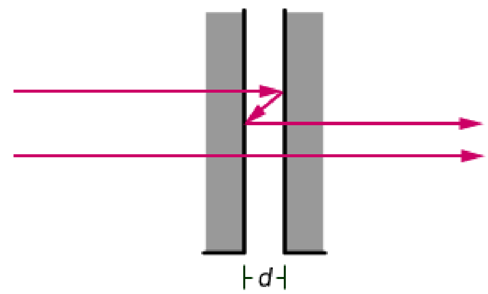
__________ m

12.2 A camera lens with an index of refraction of 1.65 is to be coated with a material with index of refraction 1.38 that will make it nonreflecting for red light whose wavelength in air is 651 nm. What is the minimum thickness of the coating?

__________ m

Additional Problems

A.1 A Fabry-Perot interferometer consists of two parallel partially-reflecting plates placed a distance *d* from each other as shown. The beam that passes straight through interferes with the beam that reflects once off each of the mirrored surfaces. (The reflected beams are essentially perpendicular to the mirrors. The angles of reflection are exaggerated − and unequal to the angles of incidence − in this diagram.) For light of wavelength 622 nm, what is the smallest, nonzero value of *d* that results in constructive interference?



__________ m

A.2 A double-slit experiment is performed underwater with a submerged HeNe laser. Suppose that the separation of the slits is 0.0142 cm, while the wavelength of the light in air is 633 nm. The screen is 2.41 m away from the slits. Find the distance along the screen, measured from the centerline, that corresponds to the second bright fringe (m = 2). The index of refraction of water is 1.33.

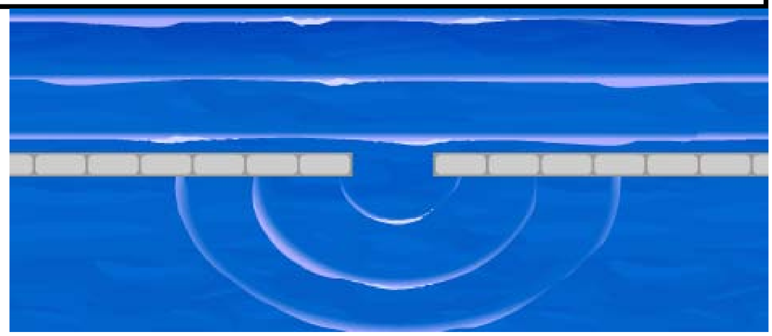
__________ m

A.3 A thin film of soap (n = 1.35) lies on top of a glass plate (n = 1.50). Visible light is incident almost normal to the plate. Maxima of reflected light are observed at two wavelengths: 412 nm and 618 nm. What is the thinnest that the film can be?

__________ m

chapter 39 Diffraction

39.0 - Introduction



Water wave diffraction.

Waves diffract, or spread out, as they pass from a narrower region to a wider one. In this chapter, we discuss the fundamental principles that cause diffraction, and analyze the interference patterns created by diffraction.

Diffraction can be observed in water waves. As a wave passes from a narrow canal into a lake, it expands. You see the expansion of a water wave in the picture above. An experimenter has created a wave whose wave front is a straight line. When the wave reaches the wall, only the portion of the wave that encounters the gap can pass through. The portion expands (diffracts) after it passes through the gap.

Sound waves also have this “spreading” ability. When you hear a friend down the hallway calling you, even though she is out of your line of sight, you hear her in part because the sound waves diffract as they emerge from the doorway. If sound moved solely in a straight line, like a bowling ball rolling straight out of her doorway, you would not hear your friend calling you. The sound would simply reflect back into her room.

For a given size opening, waves of longer wavelength diffract more than those of shorter wavelength. Because light has a much shorter wavelength than sound, it spreads out far less after it passes through the doorway. This is why you can hear the voice of your friend, even when you cannot see her.

The wavelength of visible light is in the range of hundreds of nanometers, so short that at the scale of everyday objects, it can be treated as if it moves in perfectly straight lines without spreading (diffracting). This is why it is not only convenient but also reasonable to use straight-line “rays” of light to analyze optical components.

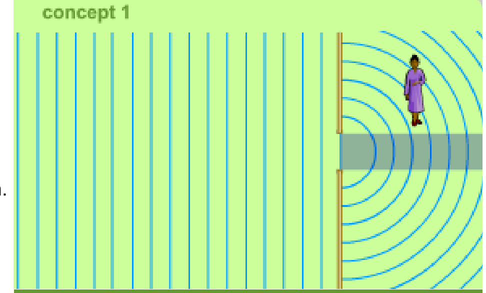
However, at smaller scales, the wavelike behavior of light cannot be ignored. Understanding this aspect of light is important to many modern applications, such as the manufacture of computer chips. Engineers who work in the semiconductor industry use visible and ultraviolet light to etch the pattern of microscopic circuits on computer chips. The light shines through thin slits and holes in an otherwise opaque “mask” to define the network of wires and other electronic elements.

As these engineers attempt to improve performance and shrink the devices by crowding the circuit components closer together, they have to deal with diffraction of the light waves as they pass through smaller and smaller apertures in the mask. To minimize diffraction effects, some are experimenting with the use of electromagnetic radiation such as x-rays, with wavelengths thousands of times shorter than the wavelengths of visible light.

39.1 - Diffraction

Diffraction: The expansion or spreading of a wave front.

concept 1



Diffraction
Wave fronts “expanding”

In Concept 1, we show the diffraction of wave fronts of a sound wave as they pass through a doorway. Only a section of each wave front can pass through the doorway. If this section of the wave front did not expand, it would move straight along the gray path. Instead, it expands spherically after passing through the doorway. The spherical expansion of the sound wave allows the person in the picture to hear the sound.

The debate over the cause of diffraction created a tale of considerable irony. The story starts with a competition sponsored by the French Academy on the subject of diffraction. The Frenchman Augustin Jean Fresnel (1788 - 1827) submitted a paper to the judges in 1818 discussing diffraction that was premised on considering light as a wave.

A noted mathematician, Simeon-Denis Poisson, had earlier mocked Fresnel’s theories, pointing out that Fresnel’s theory predicted that the shadow of a circular opaque object subjected to a bright light would exhibit a bright spot in its center. Since the circular object blocks the center portion of a screen from the light source, this result is highly surprising.

Noted mathematicians can be wrong (or should trust their mathematics). The Fresnel/Poisson debate was resolved by experiment. In Concept 2, the reddish image on the screen was created by shining a bright light at a small ball bearing. You can see the expected bright light surrounding the circumference of the ball, but − surprise! − there is also a bright spot at the center of the shadow caused by light diffracting around the ball’s perimeter.

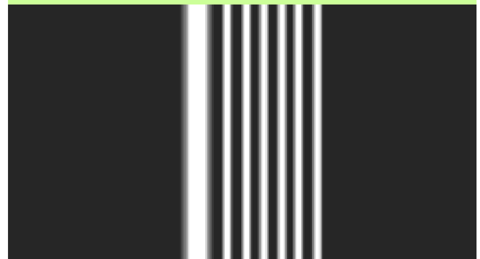
Fresnel was right, and Poisson unintentionally helped to confirm his theory. The irony is that though Fresnel did all the work, and Poisson initially ridiculed the theory, the bright center today is often called the *Poisson spot*. Others call it the *Fresnel spot*, perhaps a more appropriate name, with *Arago spot* yet another name (François Arago was the judge of the competition).

Waves also expand around sharp edges into regions that would otherwise be in shadow if the wave traveled only in a straight line. This type of diffraction causes an interference pattern. The photograph in Concept 3 shows *straight-edge diffraction*. Light from a point source passes by a sharp, well defined edge, and the resulting diffraction causes a pattern on a screen behind the object.

concept 2

Poisson spot
Bright spot caused by diffraction

concept 3



Straight-edge diffraction
Causes interference pattern

39.2 - Huygens' principle

The Dutch scientist Christian Huygens (1629 - 1695) concluded that light was a wave, made up of tiny points that emit spherical *wavelets* like the ones you see in Concept 1 to the right. Although his model has its limitations, it provided a basis for explaining many phenomena that scientists observed, and his model was usefully employed and expanded. Here we explain his fundamental principle, and show how his model can be used to explain diffraction.

As Huygens wrote:

> …each particle of matter in which a wave spreads, ought not to communicate its motion only to the next particle which is in the straight line drawn from the luminous point, but that it also imparts some of it necessarily to all the others which touch it and which oppose themselves to its movement. So it arises that around each particle there is made a wave of which that particle is the center.

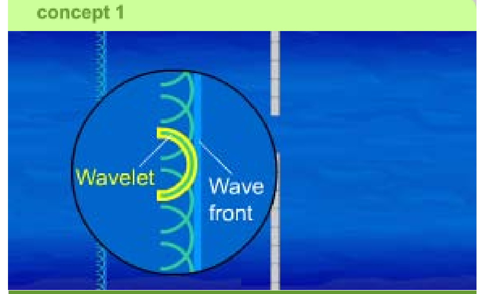
In Concept 2, you see the wave front created by the wavelets moving toward a barrier with a gap in it. Before the wavelets reach the barrier, each wavelet has another wavelet adjacent to it.

Huygens' principle states that the position of the wave front at any time can be found by drawing a line tangent to the leading surface of each wavelet. As each wavelet moves forward, the wave front moves forward as well.

Let's consider what happens when the wave front meets the barrier. The barrier prevents most of the wave front from moving forward (for simplicity's sake, we ignore the reflection of the wave). One section of the wave front (and the wavelets that make it up) passes through the opening. They continue to move forward after passing through.

As Huygens mentions in the quote above, you can think of the sources of the wavelets as a set of vibrating particles. After the wave front passes through the slit, the oscillating particles that now make up the wave front are free to interact with other particles above and below them. The interaction of the particles with their formerly stationary neighbors and the spherical form of the wavelets account for the spherical expansion of the wave front. This is illustrated in Concepts 3 and 4.

concept 1

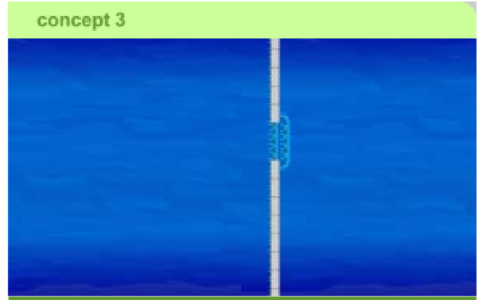


Huygens' principle and diffraction
Wave front made up of spherical wavelets

concept 2

Huygens' principle
Wave front tangent to wavelets, same speed

If you like, you can think of the wavelets as a group of subway passengers exiting a subway car. Their "wave" is constrained by the width of the subway doors, but when they exit the subway and enter the subway station, their wave can expand amongst the people awaiting the subway.



Diffraction
Portion of wave passes through slit
After the slit, particles at edges interact with nearby particles

Diffraction
Result is a spherical wave front

39.3 - Physics at work: computer chips

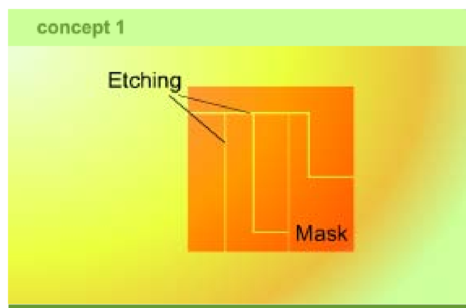
Light is used to help create the circuits on computer chips. The process employed to "draw" these microscopic circuits is called *photolithography*. The name is quite apt: it comes from Greek words that combine to mean "writing on stone with light."

In photolithography, chip manufacturers first create a *mask*, which is a stencil of the pattern of circuitry they want on the chip. A blank chip is covered with a light-sensitive polymer layer (called *photoresist*) that undergoes chemical changes when exposed to light. Short-wavelength light is then shined through the mask onto the chip. The next step is to dip the device into a developer that etches away the photoresist only where light had struck, leaving behind the pattern defined by the mask. This is then followed by further steps that deposit metal, ions, or semiconducting material into the gaps etched away by the developer. The process is repeated many times; each time, a new layer of circuitry is etched on the chip.

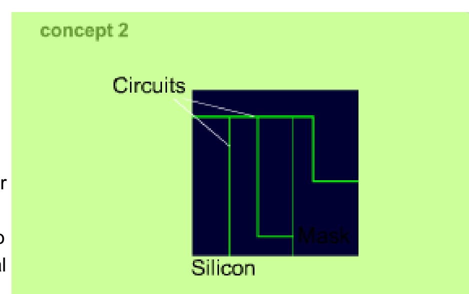
Diffraction poses an increasing problem to semiconductor manufacturers. In the 1970s, microprocessors contained less than 10,000 transistors. Today, commercial manufacturers create microprocessors that contain approximately 100 million transistors. In the future, engineers want to place even more circuits on the same surface area. *Moore's law* describes this rapid increase in the number of transistors. It states that the number of transistors on an integrated circuit doubles every 18 months. This law, stated by Intel cofounder Gordon Moore in 1965, has proven to be very accurate.

As manufacturers add more and more transistors, the circuit paths must become thinner and thinner and are squeezed closer and closer together. Circuit designers want the circuits to be as narrow and as close together as possible. Diffraction causes the light to expand after passing through the mask, and this creates wider circuits, and the potential for overlap.

To minimize the effects of diffraction, researchers are exploring new processes that use x-rays. Because x-rays have a shorter wavelength than light, they diffract less. Although the benefits of techniques such as x-ray lithography may be significant, some scientists fear that diffraction will ultimately present a fundamental limitation to circuit density.

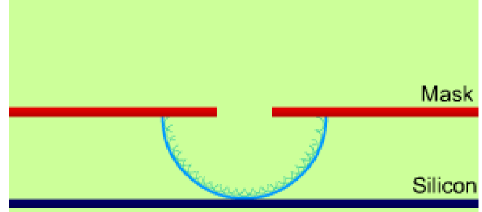


Mask
Stencil of circuit pattern



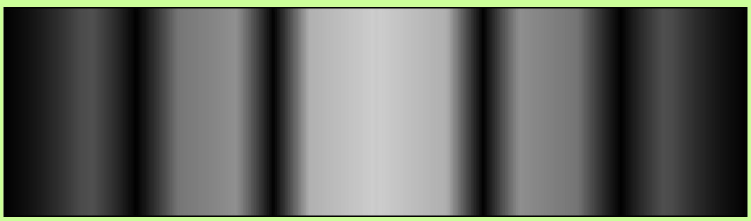
Circuit
Etched by light

concept 3



Diffraction limits circuit density
Causes “blurred” or wider circuits

39.4 - Single-slit diffraction



Single-slit pattern

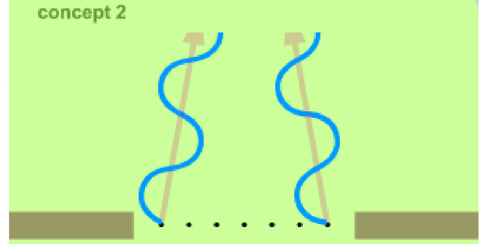
concept 1

Single-slit diffraction pattern
Light and dark fringes

Waves, including light waves, diffract as they pass through a single narrow slit. When a screen is placed on the other side of the slit, a *diffraction pattern* will be created on the screen. This pattern arises from interference among light waves coming through different portions of the slit. We will assume that the screen is far enough away from the slit that the rays that pass through are approximately parallel to each other. This is called *Fraunhofer diffraction.* (*Fresnel diffraction* occurs when the rays cannot be treated as parallel, and is discussed in more advanced texts.)

The result of Fraunhofer diffraction is a pattern of light and dark bands (often called *fringes*) on the screen, as shown above. The black-and-white image emphasizes the light and dark pattern.

concept 2

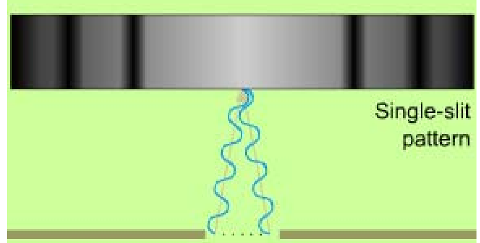


Use Huygens' principle
Model light as emanating from point sources

This pattern of light and dark can be explained using the concept of interfering waves. As with double-slit interference, bright fringes result from constructive interference of waves and dark fringes from destructive interference. The intensity of the bright fringes diminishes the farther they are from the midpoint. The first bright fringe is located straight across from the single slit. The other fringes are located in a symmetric pattern on both sides of the center.

In the illustrations, we simplify the configuration necessary to produce the diffraction pattern shown. A lens is typically used to focus the light, making the diffraction pattern clearer.

concept 3



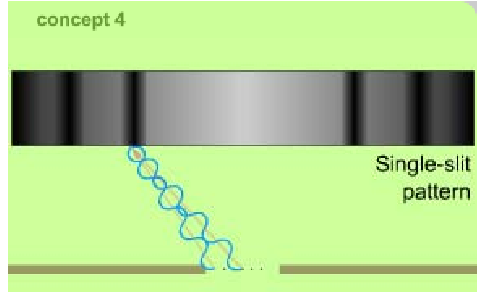
Constructive interference
Causes bright fringes

To explain the source of the interference, we use Huygens’ principle and treat the light passing through the slit as though it were made up of individual waves (wavelets) emanating from a series of point sources, as shown in Concept 2. We focus on the waves emanating from just two of those point sources to simplify the drawings and explanations. As with double-slit interference, the difference in the waves’ path lengths to the screen (and any resulting phase difference) determines whether they interfere constructively, and create a bright fringe, or destructively, to create a dark fringe.

First, we show constructive interference. In Concept 3, we consider two waves from the edges of the slit. They meet at the center of the screen. Since they travel the same distance, there is no path length difference, which means they arrive in phase. Any point source within the slit can be matched to a corresponding “mirror point” that is an equal distance from the midpoint, but on the other side. The interference is completely constructive, and the result is the central bright fringe, which is the brightest in the entire diffraction pattern.

In Concept 4, we show how completely destructive interference creates a dark fringe. In this case, we consider a wave on the left edge of the slit and a wave from the center of the slit. The wave on the left travels one-half wavelength less to reach the screen than the wave on its right. This means they will be completely out of phase when they meet at the screen. In fact, every wave has a corresponding wave exactly half a slit away that will cancel it at that screen location.

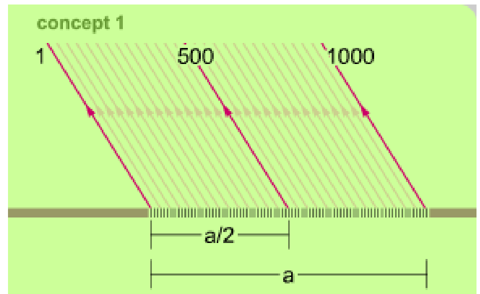
The regions between the lightest and darkest points are the result of intermediate interference. In these regions, the overall interference is neither completely destructive nor completely constructive, and the brightness at these points is between that of the points just discussed.



Destructive interference
Causes dark fringes

39.5 - Single-slit diffraction: locating the minima

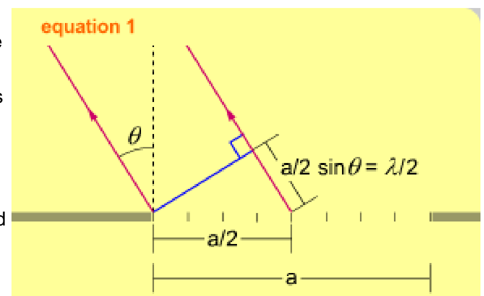
Geometry can be used to locate the dark fringes, or minima, in a single-slit diffraction pattern. Their location depends on the wavelength of the light and the width of the slit. The location is expressed in terms of the angle at the slit between a line to the center of the diffraction pattern and a line to the dark fringe. The general approach of this work will remind you of the work done to locate minima in double-slit interference patterns.



Single-slit diffraction
Destructive interference causes dark fringes

In Concept 1, we show the basic model we will use. A large number of points are the sources for parallel light waves or rays. These waves all converge at a single point, the location of the first dark fringe. Parallel lines that converge may seem unrealistic, but remember this diagram is not drawn to scale, and lines that are close to parallel would converge at sufficient distance.

Each wave is paired with a wave on the “other” side of the midpoint of the slit, half a slit width $a/2$ away. All the waves are initially in phase at the slit. At the location of the first dark fringe, there is perfect destructive interference between each wave and its partner that is $a/2$ distant, because at that angle, their path length difference to the screen is half a wavelength.



Location of minima

First minimum:

$$\sin\theta = \frac{\lambda}{a}$$

Any minimum:

$$\sin\theta = m\frac{\lambda}{a}$$

θ = angle between normal line, ray
m = 1, 2, 3....
a = width of slit

In the illustration for Equation 1, we show the figure used to derive the equation for the position of the dark fringes. With the screen located at a large distance compared to the width of the slit, the angle θ is approximately the same for all waves. The path length difference equals the leg of the right triangle shown in Equation 1. We can calculate this distance by multiplying the hypotenuse of the triangle, which is $a/2$, by $\sin\theta$. For destructive interference between each wave and its partner, this path length difference equals one half the wavelength, which results in the first equation shown to the right.

How do you locate other dark fringes? Imagine splitting the slit into four sections instead of two, each with a width of $a/4$. Partner each wave with another wave that starts out $a/4$ distant. If they are to destructively interfere as before, their path length difference must again be $\lambda/2$. Now we write that $(a/4)\sin\theta = \lambda/2$, or $\sin\theta = 2\lambda/a$. At this angle, the path length difference between pairs of waves is exactly half a wavelength. The two waves cancel at the screen, and the same can be said for any two waves that originate at points separated by $a/4$. The equation above determines the locations of the second set of dark fringes from the center.

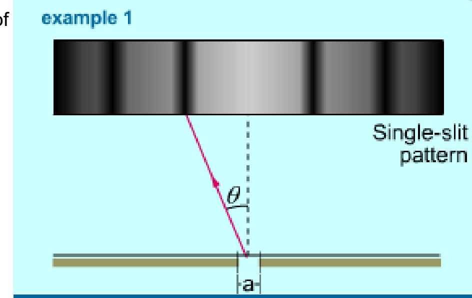
We can continue. We could divide the slit into 6 parts and partner each wave with another “canceling wave” that starts out at a point $a/6$ distant. Doing this would enable us to determine that the third set of dark fringes occurs when $\sin\theta = 3\lambda/2$. This reasoning can be continued, and leads to a general equation to locate dark fringes, which is shown to the right.

Note that we have located the dark fringes without discussing the locations of the bright fringes. The centers of the outlying bright fringes are approximately halfway between the centers of the surrounding dark fringes. For instance, you could use the locations of the dark fringes nearest the center as a way to determine the width of the central bright fringe. Its width can be defined as the distance between the centers of these fringes.

We use the dark fringe equation to further analyze a predicament that semiconductor manufacturers face. We have already discussed the problem of diffraction widening the circuits etched onto a chip. This limits how close the circuit lines can be and the number of circuits that can be put onto a certain size chip. One possible way around this is to decrease the width of the circuit lines. Unfortunately, diffraction works against this strategy and this can be shown via the dark fringe equation. Note that as the width of the slit decreases, the angle to the first dark slit increases. In other words, the center bright fringe becomes wider. A mask that creates narrower circuits suffers from increased diffraction.

One answer to the predicament can be deduced from the equation. The smaller that λ is relative to a, the smaller the angle θ. Semiconductor manufacturers already use very short wavelength light, such as ultraviolet, and are experimenting with using x-rays instead of light in order to further minimize the effects of diffraction. x-rays have a smaller wavelength than visible light, which means their central bright region is narrower.

example 1



Light with wavelength 630 nm is incident on a slit with width *a*. If *θ* = 11.0° for the first minimum, what is the slit width?

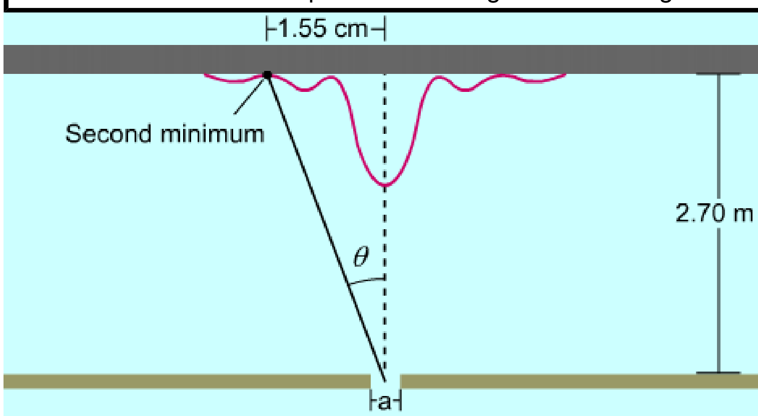
$$\sin\theta = m\frac{\lambda}{a}$$

$$a = \frac{m\lambda}{\sin\theta}$$

$$a = \frac{(1)(630\times 10^{-9}\,\mathrm{m})}{\sin(11.0°)}$$

$$a = 3.3\times 10^{-6}\,\mathrm{m}$$

39.6 - Interactive checkpoint: calculating minima in single-slit diffraction



Light from a red helium-neon laser, with wavelength equal to 633 nm, illuminates a slit that is cut in a thin foil. On a screen 2.70 meters away, the distance between the second diffraction minimum and the central maximum is 1.55 cm. Calculate the angle of diffraction of the second minimum, and the width of the slit.

Answer:

$\theta =$ ______ °

$a =$ ______ m

39.7 - Diffraction by a circular aperture

At the right is a photograph of the diffraction pattern created by light passing through a circular aperture, such as a telescope lens or the pupil of the human eye. Instead of the image being a distinct point, it consists of a central circular bright region, surrounded by alternating rings of light and dark. (The image has been overexposed to highlight the outlying bright fringes. They would be less intense than this illustration implies.)

The equation at the right is used to locate the first minimum in the diffraction pattern of a circular aperture. We state this equation without deriving it. It is analogous to the equation for locating the first minimum produced by a single slit, $\sin\theta = \lambda/a$.

One practical conclusion can be drawn from this equation: θ will be smaller when d is larger. A larger opening will suffer from less diffraction. When you next look at optical equipment, you may see how engineers and nature have factored this relationship into their designs. For instance, when you see the large size of research telescopes or the relatively large eyes in predators, you are viewing designs that limit the effects of diffraction. These designs serve other purposes as well.

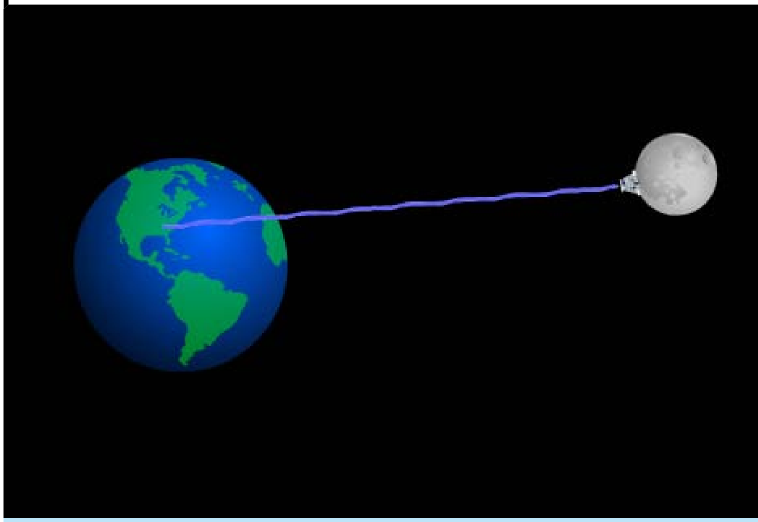
equation 1

Diffraction by a circular aperture

$$\sin\theta = 1.22\frac{\lambda}{d}$$

θ = angle to first minimum
λ = wavelength
d = diameter of aperture

39.8 - Interactive checkpoint: Dr. Evil's evil laser



Dr. Evil™ proposes using a moon-based X-ray laser to drill a narrow hole through the roof of Fort Knox, 3.85×10^5 km away, to steal…a megadollar. Luckily for law enforcement agencies, his doctorate is in evil, not in physics.
His laser operates at 2.33 nm, and emits radiation from a circular aperture of diameter 0.781 mm. (a) Find the angular width in radians of the beam's central maximum. (b) Find the diameter of the beam's central maximum at the Earth. Why will Dr. Evil™ give you the evil eye if you tell him this result?

Answer:

θ_{width} = ____ rad

D = ____ m

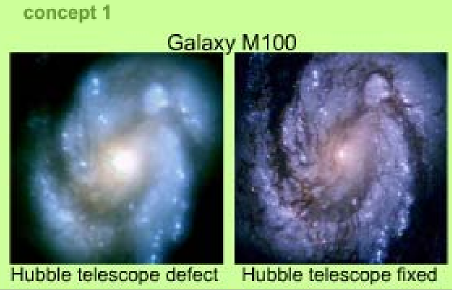
39.9 - Resolving power

Resolving power: Ability to distinguish between two objects.

Resolving power expresses the capability of an optical system to show the separation of objects that are close together. This correlates to the instrument's ability to show fine detail. A microscope is one example of an instrument that needs a good deal of resolving power. To be effective, a microscope must enable you to distinguish very small, close details. Telescopes also must supply great resolving power to allow the viewer to separate distant objects.

In concept 1 we show the importance of resolving power. The image of galaxy M100 on the left was taken by the Hubble Space Telescope before a defect in its mirror was corrected; the image on the right was taken after the defect was fixed. The right-hand image shows much more detail, and many more distinct objects are visible due to the greater resolving power achieved.

concept 1



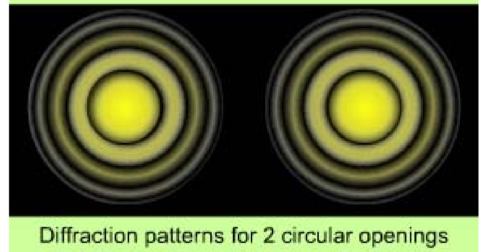
Resolving power
Measures ability to distinguish between two objects

The Hubble telescope can resolve light sources that are less than $0.00003°$ (≈ 0.0000005 rad) apart. What does that mean on a human scale? It could resolve a pair of headlights roughly 3000 km away, approximately the distance from Denver to Miami. A great conceptual design, some repairs and ongoing maintenance, and the fact that the telescope is above the blurring effects of the Earth's atmosphere, all contribute to this mind boggling capability.

You do not need a telescope to understand what is meant by resolving power. You can experience it simply by looking at a car's headlights at night. At a large distance, the car's lights look as if they come from a single source, in one blur of light. As the car comes closer, the blur resolves into two separate sources of light.

Diffraction causes the stars and the headlights to "blur" together. Concept 2 shows the diffraction pattern created by two sources that are far enough apart that their diffraction patterns do not significantly overlap. Concept 3 shows what occurs when the diffraction patterns overlap. One "blurry" central image is the result.

concept 2



Two resolved images
Diffraction patterns remain separate

concept 3

Blurry image
Significant overlap of diffraction patterns

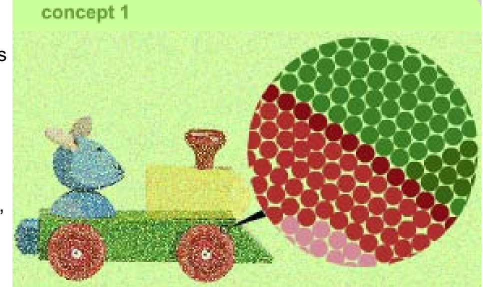
39.10 - Resolving power of the eye

Human eyes, and those of other animals, resolve objects. From ten meters away, a person with good eyesight can resolve two objects separated by only three millimeters. Some animals can do far better. At the same distance, an eagle can resolve two objects that are only $0.8\ \text{mm}$ apart. This ability allows eagles in flight to pick out potential prey on the ground with amazing accuracy. For example, an eagle can distinguish a rabbit from its surroundings from as far as a mile away.

Painters and other artists sometimes exploit the limitations of the human eye to create their effects. The style of painting known as Pointillism was famous for this. To the right, we have created our own Pointillist art. The "zoom-in" shows that the picture is composed of many small "dabs" of color. As you step back, your eye can no longer resolve the individual dabs, and you see the collective image.

If your school lacks a first rate collection of Post-Impressionistic paintings, the same effect is demonstrated in your daily newspaper. Look closely enough at a picture and you will see the dots that compose it.

concept 1



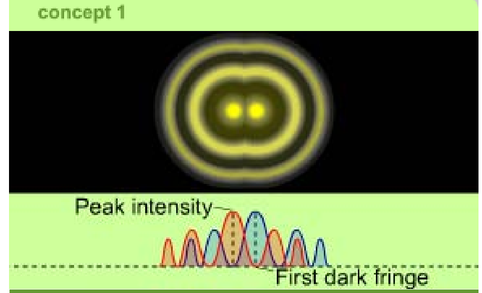
Lack of resolution leads to art

39.11 - Rayleigh's criterion

Rayleigh's criterion: Two point light sources are "just" resolved when the first dark fringe of diffraction of one intersects the central bright fringe of the other.

British scientist Lord Rayleigh established a definition for the limit of resolution, now called Rayleigh's criterion. It states that two point light sources are resolved (distinct) when the first dark fringe of one object intersects the central bright fringe of the other. You see Rayleigh's criterion applied in Concept 1 and in Equation 1. Light from two sources passes through an opening, and diffracts before it strikes a viewing screen (which could be a piece of film, or a retina). The first dark fringe of each pattern overlaps the central bright point of the other one.

concept 1



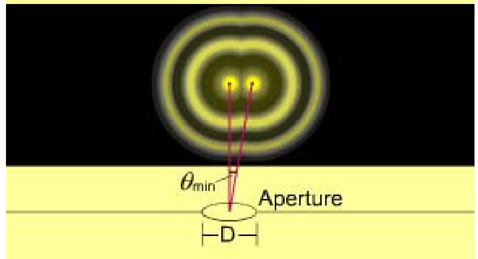
Rayleigh's criterion
Two point objects just resolved when first dark fringes intersect central bright fringes

Rayleigh's criterion is applied to openings in optical instruments, such as the lens of a telescope or the pupil of an eye. These openings are often circular, and the equation used for the angle to the first circular dark fringe is $\sin\theta = 1.22\lambda/d$, where d is the diameter of the circular aperture. For these circular fringes,

θ is measured from the center of the circular central bright fringe. This angle is shown in Equation 1. The equation lets you calculate the minimum angle between two objects that is required for the objects to be resolved. It requires an approximation that is true for small angles stated in radians.

The equation indicates why, in the domain of animal vision, a larger pupil size leads to better resolution. As D (in this case, the pupil diameter) gets larger, the smallest angle θ_{min} that can be resolved gets smaller. Animals with larger pupils can not only resolve objects separated by smaller angles, but they enjoy superior light-gathering ability, enabling them to operate better under low light conditions.

equation 1



Resolution circular aperture

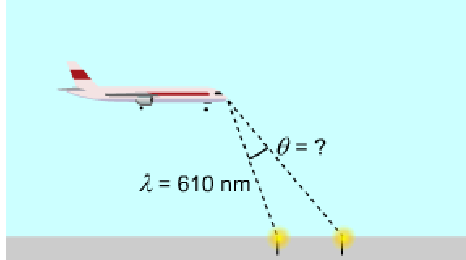
$$\theta_{\text{min}} \approx \frac{1.22\lambda}{D}$$

θ_{min} = minimum angle required for resolution (rad)

λ = wavelength

D = diameter of aperture

example 1



A pilot is looking for two bright beacons, with light of wavelength of 610 nm. The diameter of her pupil is 2.7 mm. What angle must separate the beacons for her to resolve them?

$\theta_{\text{min}} \approx 1.22\lambda/D$

$\lambda = 610 \text{ nm} = 0.61\ \mu\text{m}$

$D = 2.7 \text{ mm} = 2700\ \mu\text{m}$

$\theta_{\text{min}} \approx (1.22)(0.61) / 2700$

$\theta_{\text{min}} \approx 0.00028 \text{ rad}$

39.12 - Sample problem: resolving double stars

The angular separation between the two stars that make up the double star Albireo is 1.66×10^{-4} radians. Can the human eye resolve these two stars?

Nearly half of the stars in the galaxy are double or multiple stars, meaning they are composed of two or more stars that are gravitationally bound and rotate around each other in stable orbits. Their period of rotation can vary from a few hours to thousands of years. Diffraction effects limit the observations of these stars, and very few double stars can be resolved by the unaided human eye. One of the most colorful double stars in the sky is Albireo, which lies at the head of the constellation Cygnus. This constellation looks like a swan and Albireo is the "beak star", the second brightest star in the constellation. The two stars that make up Albireo are approximately 6.6×10^{14} meters apart, and no one knows if they are truly gravitationally bound. However, it has been calculated that if they do rotate around each other, their period is roughly 7,300 years! When looking at the two stars the brighter one appears red-orange (and is actually a double star itself), while the dimmer one is blue white.

In this problem assume you are looking at the stars under perfect conditions and that the wavelength of the light coming from them is 550 nm (about the middle of the visible spectrum). The approximate diameter of the human pupil at night is between 4 and 7 millimeters.

Variables

Minimum angular separation	$\theta_{\min} = 1.66\times10^{-4}$ radians
wavelength of incident light	$\lambda = 550\times10^{-9}$ m
diameter of aperture	D

What is the strategy?

1. Model the pupil of the eye as a circular aperture, and use the equation for the resolution of a circular aperture.
2. Solve for the diameter of the aperture required to resolve the two stars.
3. Compare the diameter of the human eye to required aperture diameter and determine whether the human eye could resolve the two stars.

Physics principles and equations

Resolution of a circular aperture

$$\theta_{\min} \approx \frac{1.22\lambda}{D}$$

Step-by-step solution

Step	Reason
1. $\theta_{\min} \approx \dfrac{1.22\lambda}{D}$	Equation for resolution of circular aperture
2. $D \approx \dfrac{1.22\lambda}{\theta_{\min}}$	solve for the diameter of aperture
3. $D \approx \dfrac{1.22\,(550\times10^{-9}\text{ m})}{1.66\times10^{-4}\text{ rad}}$	enter known values
4. $D \approx 4.04\times10^{-3}$ m	final result

Since the diameter of the pupil at night is usually larger than the diameter of aperture required (4.03 mm), the human eye should be able to resolve the two stars in Albireo. However, in reality this is not possible because of atmospheric effects, which have been ignored. With a telescope finder or a good pair of binoculars you will be able to resolve the pair. The Cygnus constellation is visible from the Northern hemisphere during the summer and early fall. Happy star gazing!

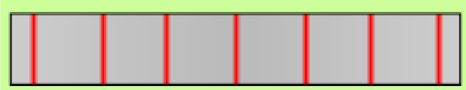
39.13 - Diffraction gratings

Diffraction grating: A material with a large number of narrow, regularly spaced slits or grooves designed to produce a crisp interference pattern.

A double-slit configuration will create a pattern of light and dark fringes. Diffraction gratings use a much larger number of slits or reflective grooves to create a very crisp interference pattern. Instead of broad fringes, the bright regions are sharp, narrow lines.

The slits in a grating are very close together and great in number. Although they are called slits, a diffraction grating often consists of a series of narrow reflective grooves, like those found in a CD. Commercial vendors sell diffraction gratings with over 10,000 grooves per millimeter. The distance between the slits or grooves is about the wavelength of light that is being diffracted. As you might suspect, manufacturing a grating with such high precision is a nontrivial task.

concept 1



Diffraction grating interference pattern
Consists of bright, narrow lines

The same basic analysis used to locate the position of dark and bright fringes with a double-slit configuration is used with a diffraction grating. The concepts of path length and phase difference, and constructive and destructive interference, again come into play. The formula in Equation 1 gives locations of the principal maxima of a diffraction grating.

There are some noticeable differences between the interference pattern produced by a diffraction grating and one produced by a double-slitted barrier. For one, the brightest fringes, called principal maxima, are much brighter and narrower for a diffraction grating. In fact, they are so crisp that they are often called lines. Between the principal maxima there are lower intensity secondary maxima. As the density and overall number of slits increases, the principal maxima become sharper and the intensity of the secondary maxima reduces to nearly zero.

An essential question is: Why does a larger number of slits produce “crisper” principal maxima than just two slits?

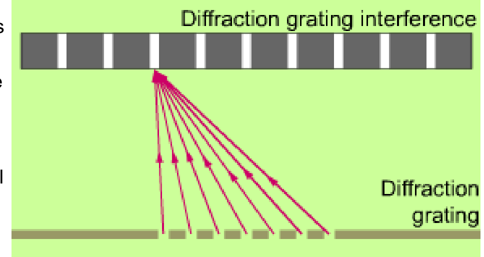
Let’s start with just two slits. For a point close to, but not exactly on a maximum, there is still a substantial amount of constructive interference, since the path length differences are minimal at this location. You have seen this in all the interference patterns you have been looking at: The fringes are far from crisp bright lines.

With multiple slits, at even a small distance from a maximum there is destructive interference. Let’s assume that at a location on the screen just to one side of the central maximum the phase difference between waves emerging from two adjacent slits is $1°$. With two slits, a phase difference of just $1°$ would mean the interference would still be quite constructive.

To discuss why this is not the case with diffraction grating, let’s number the slits in a diffraction grating sequentially: slit 1 is next to slit 2, which is next to slit 3, and so on. Again, we assume the phase difference from adjacent slits is $1°$. With the diffraction grating, the wave from slit number 181 will be $180°$ out of phase with slit number 1, and slit number 182 will be $180°$ out of phase with slit number 2, etc. Each slit can be paired with another slit that is $180°$ out of phase, and the result is significant destructive interference near the point of perfect constructive interference.

In the analysis above, we made our usual assumption that the light was monochromatic (of one wavelength). If you have a CD nearby, you have a makeshift diffraction grating, a series of closely etched slits on a reflective surface, and you can see what happens to white light. The white light reflects from the surface and, because the interference pattern varies with the wavelength of each color component of the light, it “breaks” the light into its components. The photograph in Concept 3 shows this phenomenon.

concept 2

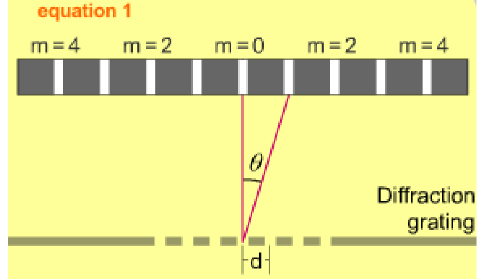


Diffraction grating interference pattern
Consists of bright, narrow lines
Caused by many, many slits

concept 3

CD acts as diffraction grating
Disperses white light into spectrum

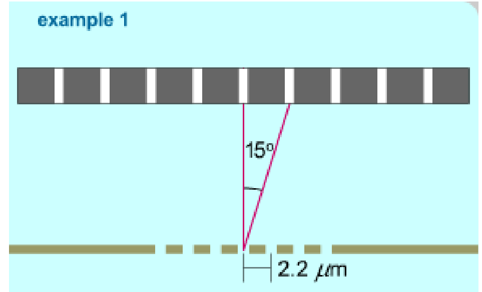
equation 1



Principal maxima

$$\sin\theta = \frac{m\lambda}{d}$$

m = 0, 1, 2, …
λ = wavelength
d = distance between slits



The first off-center maximum occurs at 15° in a diffraction grating with slits 2.2 μm apart. What is the wavelength of the light?

$\sin \theta = m\lambda/d$

$\lambda = d(\sin \theta)/m$

$m = 1$

$\lambda = (2.2\ \mu\text{m})(\sin 15°)/1$

$\lambda = 0.57\ \mu\text{m} = 570\ \text{nm}$

39.14 - Physics at play: CDs and DVDs

Most computers today contain a CD-ROM drive and many contain a DVD drive. A CD or DVD can hold a great quantity of information. The drive relies on the principles of interference to “read” the disc.

A CD or DVD contains a long spiral track with *pits* in it. These pits are formed in a disc by an injection molding process, and represent some of the smallest mechanically manufactured objects. A thin layer of metal such as silver or aluminum covers the pits. This layer in turn is covered by a thin layer of plastic.

The pits are created on the top of the CD or DVD but the disc is read using a laser that is projected up from the bottom. From the bottom of the disc, the pits appear to be raised areas. Non-pitted areas of the disc are called *land*. (An incidental fact: The pits are nearer the top of the disc than the bottom, so scratches on its top are more likely to damage the CD than scratches the bottom, or “reading” side!)

CDs or DVDs created by *burners* do not create pits in the fashion described above, but rather change the color of a layer within the disc.

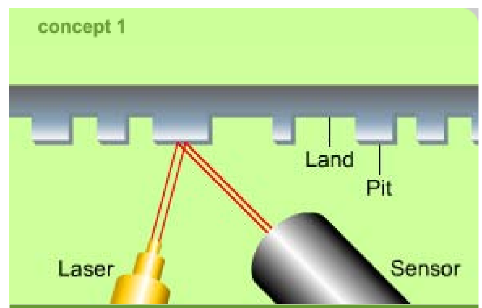
The CD or DVD reader contains a laser diode that emits a beam of light that reflects off the disc. The intensity of the reflected light varies as the disc rotates and the light reflects off pits and land. The intensity is measured and interpreted as a series of ones and zeros (digital information) by photodetectors. This information is then relayed to other systems that interpret it.

How does interference factor in? The laser beam reflects off of the CD. If all of the beam hits a land or a pit, then the path length difference back to the photodetector is essentially the same, and the result is constructive interference: bright light. You see this case in Concept 1.

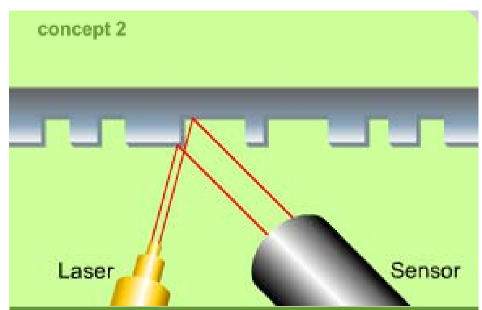
On the other hand, when the disc moves and laser light is half on a pit, and half on the land, the path length difference is significant. The two parts of the laser beam have a total path length difference of one-half a wavelength, and the result is destructive interference: darkness.

You see this in Concept 2, where we emphasize “sides” of the same laser beam, and how one side reflects off of a pit and the other off a land.

DVDs contain more data than CDs and employ a variety of strategies to do so. For instance, DVD drives use lasers with shorter wavelengths. A shorter wavelength means smaller pits are possible, and these smaller pits can be placed more closely together, allowing more data to be stored.



CDs and DVDs
Laser shines light onto track of CD, DVD
Discs have “pits” in surface
Sensor reads signal from laser



The signal
Interference pattern depends on where light strikes
Sensor receives brighter or dimmer light

39.15 - Grating spectroscope

Diffraction gratings are used in devices called grating spectroscopes to measure the wavelength of radiation. Some sources of radiation, like stars, emit light of differing wavelengths that reflect their composition. For instance, spectroscopes were used to determine that hydrogen and helium are the major components of stars.

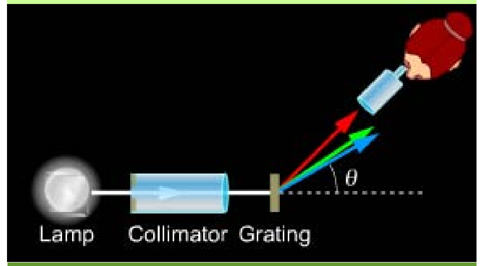
We show a grating spectroscope in Concept 1. The light first passes through a *collimator tube*, which focuses it into a tight beam and also masks off stray light. It next passes through a diffraction grating. The telescope you see the observer using can be rotated to view the maxima of the various wavelengths of light. We use the angle θ to describe the path the light follows.

As discussed earlier, the location of fringes/lines other than the central one depends on the wavelength. As the wavelength increases, so does the angle θ to a given fringe for a fixed value of m. Let's consider the $m = 1$ fringe. The wavelength of red is greater than that of blue, so the angle θ for red to that fringe is greater than the angle for blue.

In Concept 2, we show visible lines in the emission spectrum of hydrogen. These lines represent the wavelengths at which hydrogen emits light. The central point is to the left in this illustration. Note how the order of colors reflects their wavelength, with violet light (the color of shortest wavelength) being the closest to the central point, and red, with the greatest wavelength, being the farthest. If scientists saw this emission pattern from a source of unknown composition, they would deduce that the source contained hydrogen.

Diffraction gratings also can be used to analyze absorption spectra. In this case, dark spectral lines are seen against a background of a continuous spectrum. This type of light occurs when continuous-spectrum light passes through a cool gas, such as hydrogen, which absorbs specific wavelengths, leaving gaps in the final spectrum.

concept 1



Grating spectroscope
Grating separates light by wavelength

concept 2

Emission spectrum of hydrogen

Used to examine a spectrum
Wavelengths correspond to specific elements

39.16 - Physics at work: x-ray diffraction

X-ray diffraction: The diffraction and interference of x-rays passing through intermolecular gaps in a crystal; used to analyze the three-dimensional shapes of molecules.

As you have seen in the preceding sections, humans use diffraction gratings to produce interference patterns. Systems that create interference patterns also occur naturally in crystals. The pattern of constructive and destructive interference of the emerging radiation depends on the relationship of the spacing between atoms in the crystal, the wavelength of the incident radiation, and the atomic arrangement. Since the interatomic spacing is on the same order as the wavelength of x-rays (10^{-10} meters), such crystals can act as natural diffraction gratings for x-rays.

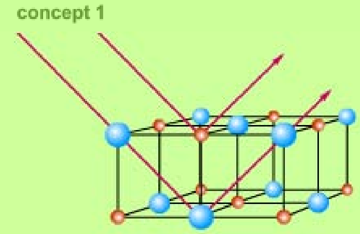
To the right is a diagram of a NaCl crystal (table salt). While most manmade interference systems, such as a diffraction grating, are designed to function in two dimensions, molecules are three-dimensional. Consequently, the interference patterns produced are far more complicated than those we have considered, but looking on the bright side (literally), these patterns can reveal much information about the crystal structure in three dimensions.

X-ray diffraction through a crystal is akin to, but not the same as, diffraction caused by slits. When x-rays enter the crystal, they are scattered: They change direction. When the rays meet on a piece of film or inside an electronic detector, they can undergo constructive or destructive interference. The process can be modeled as though the x-rays reflect off planes defined by layers of molecules. The phase difference between two waves depends upon the path length difference.

X-ray diffraction creates interference patterns like the one you see in Concept 2. Molecules as complex as DNA have been studied in this manner. In fact, this is how the famous double helix structure of the DNA molecule was discovered.

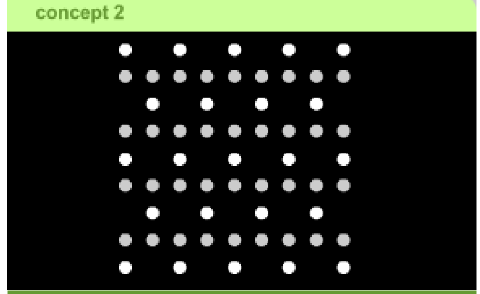
Bragg's law, shown in Equation 1, is used to analyze the results of x-ray diffraction by locating the position of the maxima as a function of the distance between planes in the crystal. Note that in this equation, θ is defined atypically, compared to other situations in optics. It is the angle between the plane defined by a layer of atoms and a ray, **not** the

concept 1



X-ray diffraction
Regular structure of crystal resembles "slits"
"Reflection" of x-rays creates interference pattern

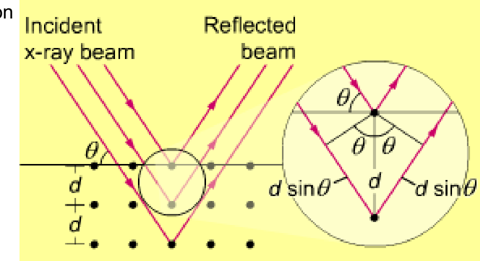
concept 2



Diffraction pattern of DNA molecule
X-ray interference pattern used to determine structure

angle between a normal line and a ray as it is elsewhere. The diagram to the right can be used to derive the equation by calculating the differences in path length as a function of the angle θ and the distance between layers of particles.

equation 1

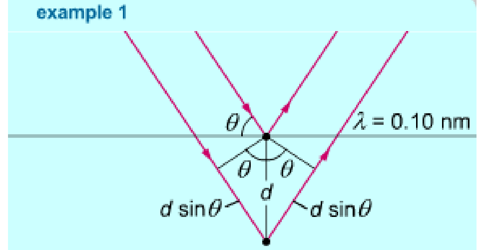


Bragg's law

$$2d \sin \theta = m\lambda$$

d = distance between planes of atoms
θ = angle between ray, plane of atoms
m = order of maximum
λ = wavelength

example 1



X-ray radiation of wavelength 0.10 nm is incident on a NaCl crystal. If the first-order maximum is at 10˚, what is the spacing between planes?

$2d \sin\theta = m\lambda$

$d = \dfrac{m\lambda}{2\sin\theta}$

$d = \dfrac{(1)(0.10 \times 10^{-9}\text{m})}{2\sin(10°)}$

$d = 0.29 \times 10^{-9}\text{m}$

39.17 - Gotchas

Diffraction only occurs with light. Not true: It occurs with sound waves, water waves, and all waves in two and three dimensions.

Single-slit diffraction is caused by interference between light rays coming from different parts of the same slit. Yes. As the light spreads out after passing through the slit, light rays from different parts of the slit travel different distances and can constructively or destructively interfere.

39.18 - Summary

Diffraction is the expansion or spreading of a wave front as it passes through an opening or past a sharp edge.

Huygens' principle is a model to explain diffraction. It says that a wave front is made up of a set of spherical wavelets.

When only one slit causes this type of interference pattern, it is called a single-slit diffraction pattern.

Dark fringes (or minima) are locations where the light destructively interferes, and bright fringes are locations of constructive interference.

The intensity of the light in a single-slit diffraction pattern is a function of the angle from the centerline.

Diffraction adds complexity to the case of double-slit interference, since each of the two slits has an associated diffraction pattern. When studying interference we ignored this effect: A graph of pattern intensity showed bright fringes of equal magnitude. When we consider double-slit interference with diffraction, the double-slit pattern is fit inside an envelope that is dictated by diffraction: The bright fringes are **not** all equal in intensity.

Diffraction also occurs with circular apertures.

Diffraction limits the ability of optical systems to distinguish between two objects, to resolve them. Rayleigh's criterion states that two objects are just resolved when their diffraction patterns overlap such that the first dark fringe of one pattern coincides with the central bright fringe of another.

A diffraction grating has a large number of slits or grooves. A grating's inference pattern has very sharp bright fringes.

Like a prism, a diffraction grating disperses polychromatic light into its constituent wavelengths. Angular dispersion is a measure of how much a grating separates two similar wavelengths of light. Dispersion increases as slit separation decreases. Resolving power is the limit of a grating's ability to produce two distinct lines for two wavelengths. Resolving power also increases as the number of slits increases.

Equations

Single-slit diffraction minima

$$\sin\theta = m\frac{\lambda}{a}, \quad m = 1, 2, 3\ldots$$

Diffraction by a circular aperture: first minimum

$$\sin\theta = 1.22\frac{\lambda}{d}$$

Resolution for a circular aperture (Rayleigh's criterion)

$$\theta_{\min} \approx \frac{1.22\lambda}{D}$$

Diffraction grating: principal maxima

$$\sin\theta = \frac{m\lambda}{d}, \quad m = 0, 1, 2, \ldots$$

Bragg's law

$$2d\sin\theta = m\lambda$$

Conceptual Problems

C.1 In a single-slit diffraction pattern, when destructive interference occurs at a dark fringe, does this mean energy has been lost from the light wave? Explain your answer.

○ Yes ○ No

C.2 State whether the angular width of the central maximum in a single-slit diffraction pattern increases, decreases, or stays the same if each of the following changes is made. Assume that each change is made independent of the others and that the apparatus is returned to its original state before the next change is made. (a) The slit width is increased. (b) The entire experiment is moved from air and instead immersed in alcohol. (c) The screen is moved closer to the slit. (d) The color of the monochromatic light used in the experiment is changed from orange to red.

(a) i. Angular width increases
ii. Angular width decreases
iii. Angular width is unchanged

(b) i. Angular width increases
ii. Angular width decreases
iii. Angular width is unchanged

(c) i. Angular width increases
ii. Angular width decreases
iii. Angular width is unchanged

(d) i. Angular width increases
ii. Angular width decreases
iii. Angular width is unchanged

C.3 Will the central bright fringe of a circular diffraction pattern be larger for blue light (450 nm) or red light (700 nm)? Explain.

i. Blue light
ii. Red light

C.4 In the chapter on interference, the equation for double-slit interference was given as $\sin \theta = m\lambda/d$. In this chapter, the equation $\sin \theta = m\lambda/a$ was presented for single-slit diffraction. In both equations, λ is the wavelength of the light, m specifies the order of a fringe, and θ is the angle to that fringe. Describe the differences between the equations.

Section Problems

Section 5 - Single-slit diffraction: locating the minima

5.1 Light of wavelength 6.50×10^{-7} m is incident on a slit with a width of 0.251 mm. A screen, placed 2.88 m from the slit, shows a pattern of bright and dark fringes. (a) Find the distance along the screen, measured from the centerline, that corresponds to the location of the first dark fringe. (b) What is the width on the screen of the central bright fringe? Reminder: The width of the central bright fringe is the distance between the centers of the adjacent dark fringes.

(a) __________ m
(b) __________ m

5.2 Coherent monochromatic light is incident on a slit whose width is 0.0211 mm. The diffraction pattern is viewed on a screen that is placed 3.12 m from the slit. The distance along the screen from the middle of the central maximum to the first dark fringe is 10.1 cm. What is the wavelength of the light?

__________ m

5.3 Light waves with a wavelength of 509 nm are incident on a long narrow aperture 0.102 mm wide. How wide will the central bright fringe be on a screen that is located 3.05 m from the slits? Reminder: The width of the central bright fringe is the distance between the centers of the adjacent dark fringes.

__________ m

5.4 A drill sergeant standing just inside a bunker doorway that is 0.951 m wide blows a whistle to summon the cadets from the exercise yard. Everyone responds except for Eli, who was about 90 m away from the bunker, at an angle of 18.0 degrees from a line normal to the doorway. Eli claims that he did not hear the whistle because he was standing at the second-order diffraction minimum for the sound waves. What whistle frequency is consistent with his explanation? The speed of sound is 343 m/s at this temperature. If you get this problem wrong, drop and give us 50.0 pushups!

__________ Hz

5.5 For a single narrow slit illuminated by red light of wavelength = 655 nm, the fourth minimum on a screen (located 11.0 m away from the slit) is 2.87 cm from the centerline. Find the wavelength of visible light for which the fifth minimum is exactly the same distance from the centerline.

__________ nm

5.6 Find the position, measured along the screen, of the first minimum for a single slit of width 5.05 μm on a screen that is 3.07 meters away, when light from a titanium-doped sapphire laser with wavelength of 702 nm is incident on the slit.

__________ m

5.7 Light of wavelength 535 nm shines through a single slit, and the diffraction pattern is observed on a screen that is placed 2.55 m away. The first dark fringes are located 1.69 cm away from the centerline on either side. What is the width of the slit?

__________ m

5.8 In an experiment, a single slit is illuminated first by light of wavelength 423 nm, and then by light of unknown wavelength. The fourth-order dark fringe resulting from the known wavelength of light falls in the same place on the screen as the third-order dark fringe from the unknown wavelength. What is the unknown wavelength?

__________ nm

5.9 Radiation of wavelength 446 nm shines on a single slit. A photodetector is placed in front of the slit, and it measures the first intensity minimum at an angle θ = 3.45° away from the centerline. What is the width of the slit?

__________ m

Section 7 - Diffraction by a circular aperture

7.1 A small circular pinhole of radius 0.125 mm is melted at the very top of an igloo. Sunlight (assume a wavelength of 5.50×10^{-7} m, which is approximately the central wavelength of visible light) streams down through the hole to the floor, which is 1.95 m below the pinhole. What is the diameter of the circular spot of sunlight on the floor? (That is, what is the diameter of the central maximum?)

__________ m

7.2 The Mystic Bright Fringes, the hippie disco band, is holding its thirty-year reunion concert at the Hollywood Bowl. The show opens with the stage dark except for a giant rotating mirrored disco ball, 2.39 m in diameter, illuminated by a laser beam shining through a pinhole aperture from the far back of the open amphitheater, 325 m away. The laser uses light of wavelength 671 nm and has an adjustable aperture. What should the diameter of the laser aperture be so that the central bright spot of the laser's diffraction pattern just covers the disco ball?

__________ m

Section 11 - Rayleigh’s criterion

11.1 A spy plane flies at a height of 21 km above Earth's surface. If you wanted to equip the plane with a camera that could resolve objects of width 1.0 cm, about enough to make out a license plate number, what diameter aperture would the camera need to have? Assume the light has a wavelength of 550 nm.

__________ m

11.2 You are playing a new game, Blackout Golf, in pitch darkness. To help you find the targets, the organizers have located two small light-emitting diodes (LEDs) on top of the flagstick at each hole. One LED is located 1.10 centimeters above the other one. You don't like playing in the dark, and your pupils are currently dilated in fear to 4.30 mm in diameter. As you contemplate a certain shot, you notice that you can just resolve the two LEDs. How far are you from the hole? (Assume the LEDs are monochromatic sources, of wavelength 6.00×10^{-7} m).

__________ m

11.3 Assume you observe the Moon under ideal conditions, both with your naked eye and by using the telescope at Mount Palomar. The diameter of your pupil at night is 4.78 mm, and the diameter of the telescope is 200 inches (5.08 m). Estimate the linear distance between two features on the Moon that can just be resolved by (a) your eye and (b) the telescope. Take the distance to the Moon to be 3.84×10^{8} m and the wavelength of light as 5.50×10^{-7} m.

(a) __________ m
(b) __________ m

11.4 The Big Dipper is probably the most distinctive asterism in the night sky, and it forms part of the constellation of Ursa Major, the Great Bear. The second star from the end of the handle of the Big Dipper is an interesting multiple star system. To the unaided eye, two stars can be seen: Mizar, and a fainter one, Alcor. They are separated by about a fifth of a degree, or 3.5 milliradians. Assume that the light coming from the pair of stars has a wavelength of 550 nm (about the middle of the visual spectrum). According to Rayleigh's criterion, what is the minimum pupil diameter required to resolve these two point sources?

(a) __________ m

Section 13 - Diffraction gratings

13.1 A diffraction grating has 980 lines/mm. What is the spacing between the centers of two adjacent slits?

__________ m

13.2 What is the highest order of interference maximum that can occur when monochromatic light with a wavelength of 560 nm illuminates a diffraction grating with 4000 lines/cm?

____

13.3 A diffraction grating with 201 lines per cm is illuminated by light of wavelength 5.50×10^{-7} m. What is the angle at which a first order maximum will be located?

__________ °

13.4 A diffraction grating has exactly 3500 lines per cm. The angle between the central maximum and the third order maximum is 35.5° for a particular diffraction pattern. What is the wavelength of the light used to create this pattern?

__________ m

Additional Problems

A.1 Light of a particular wavelength shines on a single slit of width 1.20e−4 m, creating a diffraction pattern with the first dark fringe at an angle of 0.224° from the centerline. If the same light shines instead on two even narrower slits, the interference pattern has its first dark fringe (the one adjacent to the central maximum) at the same angle. What is the distance between the centers of the two slits?

__________ m

chapter 40 Special Relativity

40.0 - Introduction

With a pair of brilliant papers published in 1905 and 1915, Albert Einstein inaugurated a revolution in physics. Scientists still are happily grappling with the implications of his work in these and other papers. His special theory of relativity and his later work predicted a range of phenomena from the amount light is “bent” by gravity, to time passing at a different rate for a passenger in a moving airplane than for an observer on the ground. To their great delight, when scientists went looking for these effects, they found them.

This revolution was all the more surprising since a distinguished scientist, Lord Kelvin, had not long before remarked: "There is nothing new to be discovered in physics now. All that remains is more and more precise measurement." Today’s physicists are more realistic, which certainly makes their research more interesting!

Strange and wonderful ideas emanate from Einstein’s work. His special theory of relativity predicts that an identical twin could leave her sister, fly off into space at nearly the speed of light and, upon her return, have aged 20 years less than her sibling. Other parts of his work, outside the realm of this chapter, have led to the discovery of black holes, objects with mass so concentrated and possessing such strong gravitational fields that even light cannot escape them.

Einstein himself found some of the implications of his research and that of his peers too incredible to believe. For instance, his work predicted that the universe is endlessly expanding. He found this idea troubling enough that he added a “cosmological constant” so the equations would predict a universe of a constant size. Subsequently, when the astronomer Edwin Hubble introduced evidence that the universe is indeed expanding, Einstein gladly removed the constant from his work, saying it was the biggest mistake he had ever made. Today, the debate about the merits of the cosmological constant continues as physicists continue their research.

Although at times seemingly “incredible,” Einstein’s essential theories have been tested and proven by scientific experiments. The results have confirmed his work to a level of precision of about one part in 10^{15}. Physicists believe that with advances in equipment and approach, they can confirm that the data conforms to his theories to even higher levels of precision.

This chapter focuses on Einstein’s special theory of relativity, as opposed to his later and more complex general theory. Einstein’s special theory of relativity is based on two postulates.

The first states that the laws of physics are the same in any inertial reference frame. What constitutes an inertial reference frame merits more discussion; briefly, you can consider any system moving at a constant velocity to be an inertial reference frame. In such a reference frame, the “classical” or Newtonian laws of physics hold true. Let’s say you are either in a train moving at a constant 125 km/h, or standing on the ground. In either situation, you can throw a ball up in the air, and predict where it will land. (Note: the Earth itself is not truly an inertial reference frame because its rotational motion means that supposedly “fixed” objects are actually accelerating. However, we typically ignore this because of its minor impact.)

In sum, the laws of physics hold true in the train and on the ground, in France as well as in Germany. This is a postulate you likely assumed: that the physics you study do not vary by location. They hold true for any inertial reference frame. You could not use Newtonian mechanics in a bumpy truck as it drove along a winding country road, but this would be a quite atypical location for you to conduct lab experiments.

Einstein’s second postulate states that the speed of light in a vacuum is the same in all inertial reference frames: It does not change due to the motion of the source of the light or the motion of the person observing the light. This insight is surprising, and does not accord with observations of everyday events.

For instance, if someone in a train moving toward you at a high speed throws a tennis ball out a window, you expect the tennis ball to be moving toward you. You intuitively add the velocity of the train to the velocity with which the ball was thrown to determine its overall velocity. You combine the velocities.

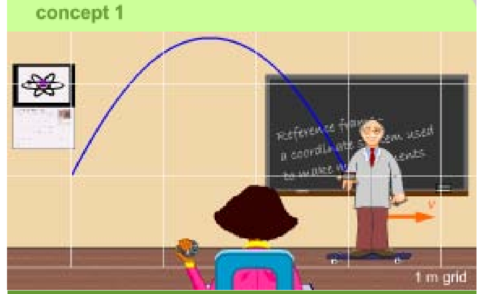
Einstein correctly stated that this is not true with light: Its speed does not change. If someone flashes a light at you from the train, whether the train moves toward you or away from you, the speed of light you measure **remains the same**.

With these two postulates, Einstein started a revolution. This chapter covers these two postulates in more depth, and then summarizes many of their amazing implications.

40.1 - Reference frames

Reference frame: A coordinate system used to make observations.

concept 1



Reference frame
Point of view for observing motion

We discussed reference frames many chapters earlier, in the context of motion in multiple dimensions. Since that was many chapters ago, we reprise the section here, though we have changed the examples and some of the discussion in order to make this section more appropriate for this chapter. Reference frames are a crucial element of special relativity.

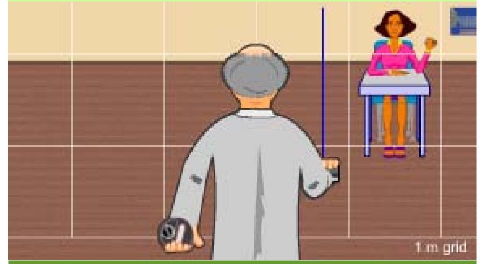
A reference frame is a coordinate system used to make observations. If you stand next to a lab table and hold out a meter stick, you have established a reference frame for making observations. Your choice of a reference frame determines your perception of motion.

We use the classroom shown on the right to discuss reference frames. The classroom shows a professor on a skateboard and Katherine, a student in the professor's class. The professor conducts a demonstration in the class: He throws an eraser up, and catches it in the same hand. The professor does all this while moving across the classroom on a skateboard. Both the professor and Katherine have stopwatches to measure the time interval between "the throw" and "the catch".

In Concept 1, we show what Katherine observes. She sees the professor moving by at a speed v. The blue arc shows the path of the eraser from her reference frame. It moves both vertically and horizontally in projectile motion. A grid is shown in the illustration. Each side of the grid is 1.0 meter. Katherine sets the position of the "throw" at $(x = 0, y = 0)$ m, and the catch at (3.00, 0) m. The eraser reaches its peak at (1.50, 1.50) m in her reference frame.

Katherine can also establish the time coordinates of these events. She starts her stopwatch when the professor throws the ball, so the throw is at $t = 0$ s. He catches the ball at $t = 1.11$ seconds.

concept 2



Reference frame
Motion is perceived relative to a reference frame

In Concept 2, we show the exact same series of events as observed in the professor's reference frame. He considers himself as stationary and views the class as moving by. It is the same experience as looking out the window of an airplane: You consider yourself stationary as you sit in a seat, and the ground is passing by.

In his reference frame, the eraser travels solely vertically. Its initial and final positions are (0, 0) m. Its peak position is (0, 1.50) m. He also has a stopwatch, which he also starts when he tosses the eraser. His observations of the time (to the precision of this stopwatch) are identical. He throws the ball at $t = 0.00$ s, and catches it at $t = 1.11$ s. To describe the catch, he could describe its space time coordinates as (0 m, 0 m, 0 m, 1.11 s). The first three coordinates state its x, y and z position, and the last states the time.

The conclusion of all this: Reference frames determine the observations made by observers. Katherine observes the eraser moving horizontally as well as vertically; the professor sees it move only vertically.

You may object: But does the professor not know he is moving? Should he not factor in his motion? Consider throwing an eraser up and down and catching it. Did you catch it at (roughly) the same position? In your reference frame, perhaps your classroom, the answer is "yes". But to an observer watching from the Moon, the answer is no, since the Earth is moving relative to the Moon.

There is no correct inertial reference frame. Katherine cannot say her reference frame is better than the reference frame used by the professor. Measurements of position, time and other values made by either observer are equally valid.

In your earlier studies, you were asked to assume that the time intervals in each reference frame were identical. At the speed the professor is moving, and to the hundredth of the second, the two observers measure the same time interval. However, if the professor were moving at, say, 75% of the speed of light, the time intervals would be quite distinct. Discussing how measured time intervals change as the observers' relative speeds approach the speed of light is a major topic in this chapter.

40.2 - Events and observers

Event: Something that can be pinpointed using position coordinates and time.

Observer: Person who records where and when an event occurs in a particular reference frame.

The definition of an event may accord with your own sense of the word − it is something that occurs at a specific place and time. A bat striking a ball is an event; the ball striking the glove of a fan in the bleachers is another event. In physics, a "Saturday night dance" is **not** an event, since it does not occur at a specific time or a specific enough location.

In the illustrations to the right, we consider a single event: The professor on the train kicks a soccer ball. He is standing on a train that is moving at a constant velocity on the track. Another observer, Sara, stands on the ground and observes the same event, the professor kicking the soccer ball. An observer is someone who records when and where something occurs in a reference frame.

Both agree that the professor kicked the soccer ball, and that it rolls along the train. But they use different reference frames to describe where and when the soccer ball was kicked.

In Concept 2, you see the position and time of the event described using Sara's reference frame. She uses a set of position markers on the ground. (You can see markers like this, though at a scale greater than meters, on railroad tracks and highways). In her reference frame, the ball was kicked at the position 2.8 meters, and her watch tells her it was kicked at 5:00 PM Pacific Standard Time.

In Concept 3, we show the same event, but now the professor describes it using his reference frame. He measures position using a scale on the train. Using that reference frame, he states that he kicked the ball at the position 0.7 meters. He also has not

concept 1

An event
"Something that happens at a specific time and location"
Example: Professor's foot meets ball

concept 2

An observer

changed his watch since he set out on his journey, so his watch tells him that he kicked the ball at 8:00 PM Eastern Standard Time.

These two sets of observations can be reconciled, but they illustrate how the reference frames of the two observers determine the coordinates of the observations they make. Both sets of observations are equally valid, given the reference frames of the observers, but they do differ. They provide a starting point in special relativity: Reference frames play an essential role in the observations made by observers.

Person who records where, when event occurs
Uses reference frame to make measurements
Sara: Foot met ball at 2.8 meter marker at 5:00 P.M. PST

concept 3

Observer in different reference frame
Observers in different reference frames make different observations
Professor: Foot met ball at position 0.7 meters at 8:00 P.M. EST

40.3 - Light can travel through a vacuum

As the 20th century dawned, the topic of how light travels generated much debate in the physics community. The evidence brought to bear on this debate helped Einstein to confirm one of his central insights: The speed of light in a vacuum is constant. The speed of light does not depend on the motion of its source or of the observer. This conclusion is one of Einstein's two postulates; the experiments that enabled him to deduce this are an interesting story in the history of science.

The primary experiment that propelled Einstein toward his conclusion was not intended to lead to conclusions about the speed of light. Rather, it was an experiment about the existence and nature of the medium through which light travels. Earlier observations had convinced scientists that light acted as a wave. Experiments had shown that a beam of light spreads out (diffracts) in the same manner as a wave of water, and that light produces interference patterns that are consistent with the patterns caused by waves as well.

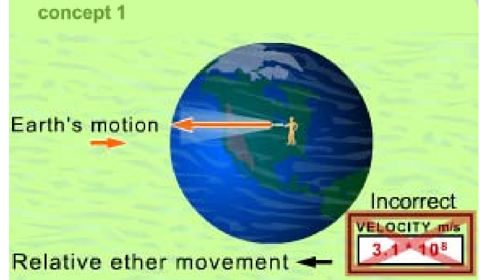
Since physicists knew light acted as a wave, they went in search of the medium in which it traveled. They reasoned that waves always move through some form of medium − water, the wire of a Slinky®, the strings of a violin. However, the medium for light was mysterious. Scientists were puzzled by the fact that light travels through the near vacuum of space, where there apparently is no medium. Physicists assumed there must be a medium, and called this elusive medium the *ether*. They assumed the ether permeated the universe, including the Earth's atmosphere.

Some physicists were skeptical about the existence of the ether. Nonetheless, they looked for ways to measure its attributes. Two clever American physicists, Albert Michelson and Edward Morley, conducted an experiment that played a key role in proving that ether did **not** exist.

Michelson and Morley reasoned: If light travels in a medium, it should move faster when traveling in the same direction as that medium and slower when traveling against it. In other words, light should act like a swimmer in the ocean: An observer on the ground would see the swimmer moving faster when swimming with the current and slower against it.

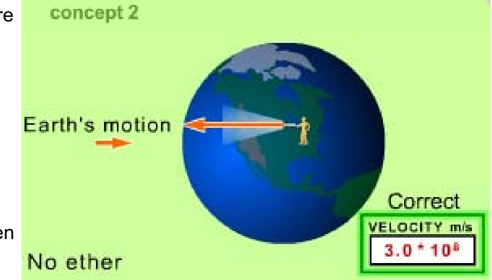
The most prevalent theory of ether stated that it was stationary and the Earth moved through it. An object stationary on the Earth would be moving through the ether. Michelson and Morley decided to measure the speed of light when the light was moving in different directions relative to the Earth's motion around the Sun. If ether existed, the speed of light would be different, and would be slowest when it moved in the same direction as the Earth's motion through the ether.

concept 1



Light travels through vacuum
If light traveled through ether
· It would travel faster when moving with the ether

concept 2



Light travels through vacuum
No difference in speed measured
· There is no ether. The speed of light is constant

Here is an analogy that may make this clearer. Imagine two birds flying at the same speed in opposite directions, while you observe them from a slow-moving train, with the train's motion causing you to feel air in your face. If the train is moving in the same direction as one of the birds, that bird will appear to fly slower than the other bird. The birds are like light, the air is like the ether, and the train is like the Earth.

The two physicists searched for changes in the speed of light using an apparatus now known as a Michelson interferometer. Michelson estimated that the Earth's motion through the ether would cause a change in the speed of light on the order of 30,000 m/s. The interferometer was capable of detecting this 0.01% change in the speed of light.

Michelson made multiple observations, measuring the speed of light with different orientations relative to the Earth's rotation and motion around the Sun. He observed no changes in its speed. As he tersely summarized: "The result of the hypothesis of a stationary ether [with respect to the Earth] is thus shown to be incorrect." (Michelson also considered the case that the ether was dragged along with the Earth; this possibility was later shown to be false as well.)

Michelson's conclusion was correct. His experiment indicated that the effects expected when waves propagate through a moving medium did **not** occur. Despite some efforts to revive the theory of ether, the clarity and simplicity of Michelson's experiment was a major and conclusive blow to the theory of ether.

If this experiment had proven only that light could travel in a vacuum, disproving the existence of the ether, it would have been a landmark in the history of science. But Einstein realized that something even more interesting was being shown by the experiment: The motion of the equipment was not affecting the observed speed of light in the way classical physics predicted it should. As the light moved in the interferometer, the equipment itself was moving, which according to the physics as Einstein had learned it should alter the effects of the experiment. In other words, even without an ether, the equipment itself was moving, and that should have affected the measurement of the light.

In his 1905 paper on special relativity, Einstein wrote, "…the unsuccessful attempts to discover any motion of the Earth relative to the 'light medium,' suggest that the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest." Michelson's experiment helped Einstein to formulate his second postulate: The speed of light is absolute; it does not depend on the reference frame. It also supported Einstein's first postulate. If the ether frame existed, the laws of light propagation would have been different in that frame, which would have contradicted the requirement that the forms of the laws of physics be the same for all observers in inertial reference frames.

40.4 - Speed of light: absolute

Speed of light in vacuum (c): 299,792,458 meters per second. Period.

One of Einstein's two postulates in his special theory of relativity is that the speed of light is **not** relative, but is the same in any reference frame. The motion of its source, or of an observer, does not change the speed.

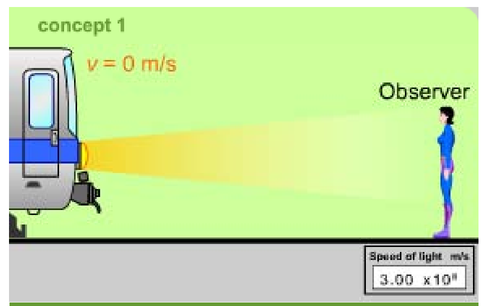
This stands in contrast with the usual principle that velocities can be added. For instance, imagine you are standing on the ground and someone in a moving car throws you a tennis ball. If the car is moving toward you, the velocities of the car and the ball have the same sign and adding them yields an increased speed for the ball. The opposite is true if the car is moving away from you.

Einstein stated that this is not true for light. In a vacuum, light travels at 299,792,458 m/s (we will typically state this with three significant digits, as 3.00×10^8 m/s). Let's say you measure the speed of the light coming from the headlights of a car. Whether the car moves toward you or away from you, your measurement is the same. The speed of light is the same whether the car stands still relative to you, is moving toward you, or is moving away.

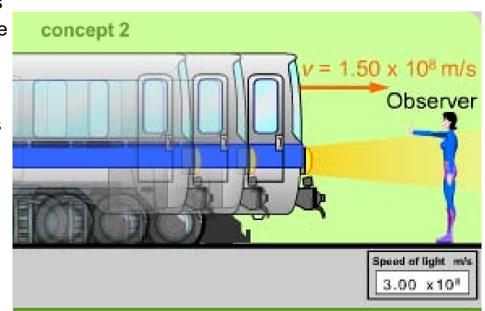
As an aside, the speed of light does differ as it moves through different media. It travels faster in air than in water, for example. But this has nothing to do with the motion of the observer and source, only the effect of the medium on the light that is passing through.

Let us say you do not believe Einstein and plan to prove him wrong. You have your eyes on a rocket that can reach a top speed of 200,000,000 m/s. You buy it and fly it to the Sun. As you are making the trip, you measure the speed of light arriving from the Sun. Then you turn around and fly back toward Earth, this time traveling in the same direction as the light. Again, you measure the speed of light. You expect the speeds to be different, but to your surprise, your instruments indicate the speed of the light you measured is the same in both directions of your travel. The velocity of your craft does not alter the measured speed of light. Whether you travel in the same direction as the light, or the opposite direction, you measure the same value.

Mind-warping, eh?

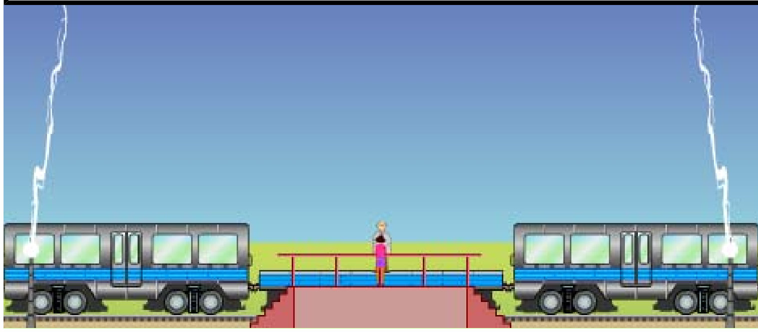


Speed of light: constant
Speed of light in a vacuum
· Precisely 299,792,458 m/s



Speed of light: constant
Motion of source, observer does not change speed of light

40.5 - Simultaneity, or the lack thereof



concept 1

Simultaneity
Events perceived as simultaneous in one reference frame
· Not simultaneous in another

Einstein's simultaneity thought experiment: The relative motion of two observers determines whether they perceive two events as simultaneous.

One of the consequences of Einstein's postulates is that the perception of "simultaneous" events is relative to an observer's reference frame. Einstein demonstrated that his postulates proved that two observers would disagree on how much time elapsed between two events, or on whether two events occurred simultaneously. In this, he disagreed with Sir Isaac Newton. Newton had stated: "Absolute, true and mathematical time, of itself, and from its own nature, flows equably without relation to anything external." Einstein was to prove this wrong.

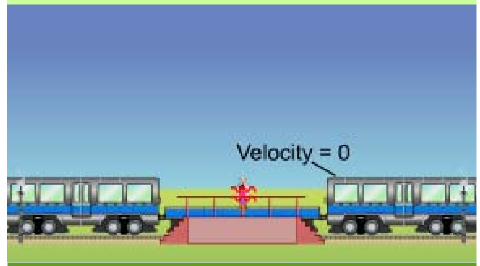
To make his point, he used a specific means to define simultaneity, noting that if an observer sees two events that are the same distance away occurring at the same moment in time, she thinks the two events occurred at the same moment in time. It seems a commonsense conclusion that another observer, also equidistant from the events, would see them occurring simultaneously. Einstein's genius lay in challenging this conclusion. He created a scenario showing that the relative motion of two observers influenced whether they perceived two events as occurring simultaneously, or as occurring one after the other.

Einstein made his case in part with a famous "*gedanken* experiment," or "thought experiment." A thought experiment is one that is conducted in the mind, as opposed to in a laboratory. Einstein used his thought experiment to make the following revolutionary point: *The measured length of time intervals differs when observers are in motion relative to one another*.

To state Einstein's thought experiment, we use a train and a platform beside the train. A professor stands still on an open railcar at the middle of the train. Another observer, Katherine, stands still at the middle of the platform. Two lightning rods, shown as gray towers in Concept 2 and Concept 3, are equidistant from Katherine and the midpoint of the platform.

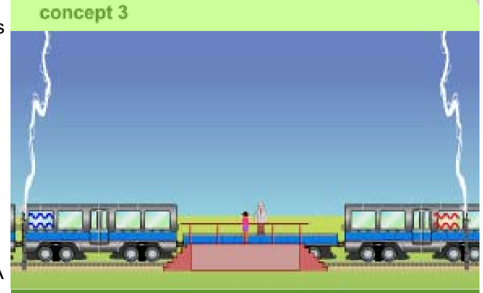
We start this experiment with the train not moving, and the professor directly across from Katherine, so he is also equidistant from the lightning rods. A lightning bolt strikes each lightning rod. These lightning strikes cause flashes of light that will be visible to the observers. Photons − packets of light − move from each lightning rod toward them. Each lightning strike counts as one of the two spatially separated events we are analyzing. We draw the photons from one rod in blue, and the photons from the other rod in red, so that you can more easily distinguish them.

concept 2



Train stationary relative to observers
Light reaches each observer at an instant in time
Both think the strikes occurred simultaneously

concept 3



Train moving relative to one observer
Professor on train: Strikes NOT simultaneous
Katherine on platform: Strikes simultaneous
Observers disagree on whether events are simultaneous

When will an observer conclude that the two events are simultaneous? The professor concludes that the bolts struck simultaneously **only** if two photons reach him at exactly the same moment. Why? Since both travel the same distance, the events are equidistant from him. The professor reasons that if the photons reach him simultaneously after traveling the same distance at the same speed, they must have started their journey simultaneously.

Katherine also concludes that the bolts struck simultaneously **only** if two photons reach her at a single instant in time. She too believes the events occurred simultaneously if two photons, traveling the same distance and the same speed, reach her simultaneously.

All this is "as expected". Two lightning bolts strike rods, photons from the strike reach each observer simultaneously, and they conclude that the two bolts struck simultaneously because the photons traveled the same distance to each observer at the same speed. This situation is shown in Concept 2, as the photons reach Katherine and the professor.

Now, Einstein changed the scenario: He put the train in motion. Einstein asked again: would they both say that the lightning bolts struck simultaneously? That the time interval between the events was zero?

The answer, as illustrated by Concept 3 is: No. In the scenario we show, Katherine still states that the lightning bolts struck simultaneously because both photons reach her simultaneously. However, as you can see, the photons from the strike on the right have reached the professor, but those on the left have not. His conclusion is that the strike on the right occurred before the strike on the left.

Although they cannot be readily seen in the illustration, this time we have the strikes make slight scorch marks on the train. The professor can later walk up and down the train and reassure himself that the distance between each strike and his position as an observer is the same.

Einstein's thought experiment with the train showed that the perception of the time interval between events depends on the observer's frame of reference. Specifically, he showed how two events that are perceived as occurring simultaneously in one reference frame would be perceived as occurring at two different moments in time (non-simultaneously) in another reference frame that is in motion with respect to the first.

Both observers were correct, based on observations in their own reference frames. This phenomenon is called the **relativity of simultaneity**. It means that time intervals between events − something that Newton assumed was absolute − are actually relative, that is, they vary depending on an observer's state of motion.

As Einstein wrote: "Are two events (e.g. the two strokes of lightning A and B) which are simultaneous *with reference to the railway embankment* also simultaneous *relatively to the train*? We shall show directly that the answer must be negative."

An observer's measurement of time intervals depends on his or her frame of reference. Time is relative. Motion alters time. This is a strange conclusion, but it is true, and it changes our view of the universe. Few thought about it before Einstein, because in his day, the effect was inaccessible to measurements at the low speeds humans were accustomed to observing. For instance, for a jogger moving at several meters per second past a stationary observer, the difference is negligible (about one part in 10^{16}), but the effect becomes much more significant as the relative speed between two observers increases.

Since Einstein's time, technological advances have made it necessary to factor this effect into the design of certain systems. You may have seen or used a Global Positioning System (GPS) unit, which receives signals from a network of orbiting satellites to pinpoint its location on the Earth's surface. The lightning strikes we considered are analogous to the signals that are emitted by satellites. The GPS unit works by interpreting the time delays in the electromagnetic signals that are received from these fast-moving satellites whose positions are precisely known. Relativistic considerations play a central role in making the system work.

40.6 - Spacetime diagrams

Spacetime diagrams are used in relativity to explain and to analyze various phenomena. In this section, we use them to explain again Einstein's thought experiment.

At the right are two spacetime diagrams used to depict essentially the same two scenarios as the prior section. Time is measured on the vertical axis in these diagrams. The horizontal axis is used to indicate an object's position as perceived by Katherine, the observer on the platform. (Be careful: This configuration is the opposite of what is typically used in the study of motion, where position is plotted on the vertical axis and time on the horizontal.)

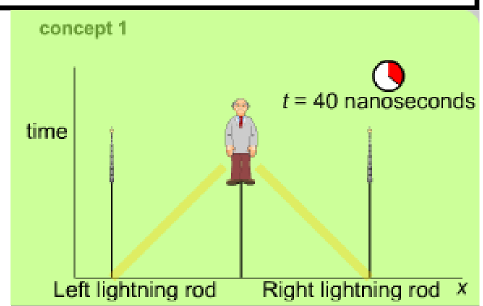
We use this diagram as a tool to discuss simultaneity, with a focus on the professor, the observer on the train.

In the first scenario, the train and its passenger, the professor, are stationary relative to the ground and the lightning rods which are firmly planted in the ground.

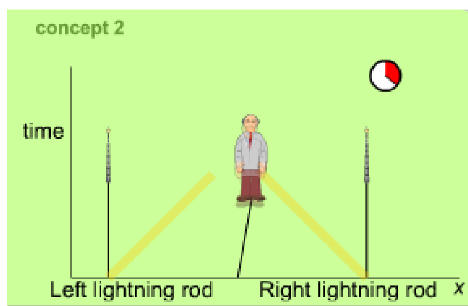
The professor's position and those of the lightning rods are shown over time. The lines representing their position over time are called *world lines*. All three are vertical lines − their position does not change over time.

The world lines of two pulses of light are shown as well. These are the flashes that emanate from the positions of the rods due to the lightning strikes. At $t = 0$, the events occur and the flashes of light from the two events start to move to the professor. Both flashes of light move toward him as shown by the yellow lines, and both reach him after 50 nanoseconds. The spacetime diagram shows the world lines of the photons in the flash. The slopes of these lines are functions of the speed of light (for the left line, c equals the inverse of the slope of the line).

In Concept 2, we show a spacetime diagram for the scenario where the train is moving. Now the professor's position changes over time, so his world line is no longer strictly vertical. (Again, the observations are made by Katherine.) Since the lightning rods remain stationary, their world lines are vertical as before. The slopes of the lines depicting the paths of the light pulses through spacetime do not change in this scenario since the speed of light is absolute.



Train stationary (relative to ground)



Train and professor moving

This means the professor intersects the world line of one flash of light before the other, and perceives the light flashes as not occurring simultaneously. You can see that he intersects the light from the right earlier in time, and at a different position, than where he intersects the light from the left. That is, the two points on the spacetime diagram are different both vertically (time) and horizontally (position).

40.7 - Interactive problem: Conduct Einstein's simultaneity experiment

There are two simulations to the right. In the first, the professor, Katherine and the train are all stationary. In the second, the professor and the train move.

As discussed before, lightning bolts strike each lightning rod. These lightning strikes cause flashes of light that will be visible to the observers. Photons − packets of light − move from the position of each lightning rod toward them. Each lightning strike counts as one of the two spatially separated events we are analyzing. We draw the photons from one rod in blue, and the photons from the other rod in red, so that you can easily distinguish them.

The lightning bolts also scorch the sides of the train as they hit the rods. The professor can use these scorch marks to confirm that the two events are equidistant from the train's center where he is standing.

Press GO to launch the lightning strikes. The simulations run in extreme slow motion. We have slowed time down by a factor of more than fifty million to clearly show what is happening.

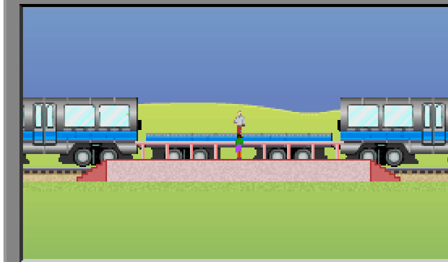
The first simulation is in Katherine's reference frame. The second simulation starts in Katherine's reference frame. You can view the same events in the professor's reference frame by pressing the "Professor's reference frame" tab.

Try each simulation. Do Katherine and the professor observe the lightning bolts as striking the lightning rods simultaneously when the train is stationary? When the train is moving?

In order to let you experiment further with this thought experiment, both simulations contain a feature on the control panel that allows you to adjust the interval of time that Katherine observes between the lightning strikes. You can use this controller to see how she and the professor observe the events as this interval changes. For instance, you can cause there to be a five-nanosecond interval of time between when one photon reaches her and the next photon reaches her. This lets you go beyond Einstein's thought experiment, where the time interval in Katherine's reference frame was always zero and she observed the lightning strikes as occurring simultaneously.

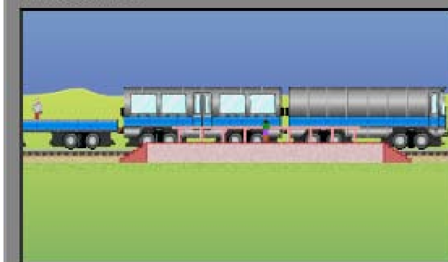
In the second simulation, can you adjust the time interval in Katherine's reference frame to cause the photons to arrive at the same instant at the professor's location? If you cause this to happen, what does Katherine now observe? Can you cause them to both believe the strikes occurred simultaneously when the train is moving? (Hint: Do not spend too much time trying!).

interactive 1



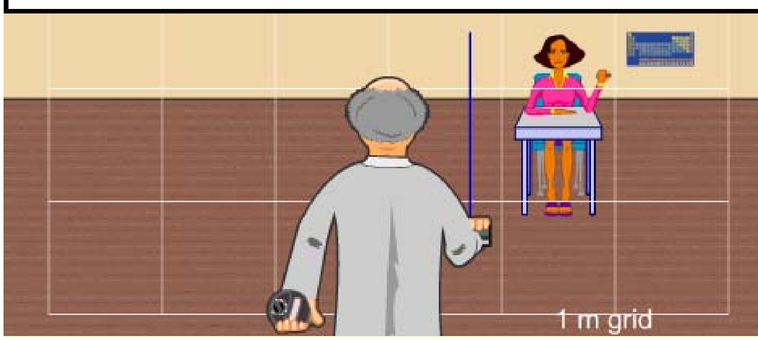
Train stationary

interactive 2



Train moves past Katherine

40.8 - Time dilation



concept 1

Proper time
Clock is at same location as events

Time dilation: A clock moving relative to an observer runs more slowly, measuring longer time intervals, than clocks at rest relative to the observer.

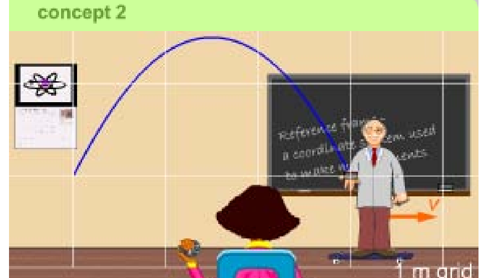
Einstein pondered: If the speed of light is constant, what does this imply? He concluded that it means that things we think of as absolute − like the time interval between events − must be "relative."

Consider the professor and the student, Katherine, in the diagrams above and on the right. The professor moves rapidly through the classroom on his skateboard. He throws and catches the ball, using his stopwatch to measure the time interval between the throw and the catch. You see this illustrated in Concept 1.

Katherine watches all this and measures the time between the throw and the catch using her stopwatch. This is shown in Concept 2.

If they compare their observations afterwards and had incredibly precise equipment,

concept 2



Time dilation
Observers differ on how much time passes between events
"Moving" observer/clock measure less time

Katherine will discover that she measured more time passing between the two events than the professor does with his stopwatch. If the professor moves at everyday speeds as we show above, the effect will be very small, far too small to be measured on any typical stopwatch. But if the professor were moving past at 87% of the speed of the light, Katherine would measure over twice as much time between the two events as the professor does.

The professor's time is said to be "dilated", which means it is passing more slowly. This is not due to some malfunctioning nor change in functioning of his clock. Any clock in his frame, including the professor's "biological clock", will record less time passing.

Time dilation can be quantified, as the equation at the right shows. To understand the equation you must understand another term: *proper time*. The proper time between two events is the time measured by a clock in the reference frame where the two events occur in the same place. "Proper" comes from the German for the events' "own time". The proper time is measured with a single clock, which is at rest in the reference frame in which the events occur at the same xyz coordinates. The professor's stopwatch measures the proper time.

Another observer, the student Katherine in this case, watches the professor (and his stopwatch) pass by at a velocity v. On the right side of the equation is the proper time interval, t_0. Katherine measures an interval of time t between the events.

We also present the equation in another way. The term that multiplies the proper time, that is, the reciprocal of the square root term, is frequently used in relativity. It is represented by the Greek letter γ (gamma) and is called the *Lorentz factor*.

Physicists have confirmed the principle of time dilation by many methods, such as studying muons. *Muons* are subatomic particles that have an average *half life* of about 2.2 microseconds. That is, in a sample of muons that are stationary relative to an observer, half of them will decay (change) into other particles during this interval of time.

Fast-moving muons are produced when cosmic rays enter the Earth's atmosphere and collide at high speeds with atoms there. The muons travel toward the ground at about $0.999c$. Scientists observe that these moving muons decay more slowly than stationary muons.

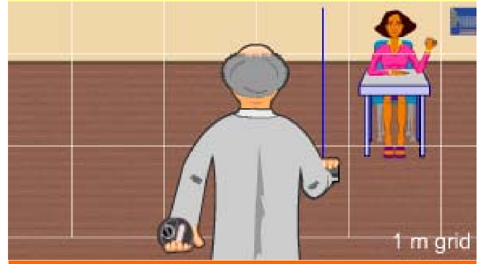
Why? The fraction of muons that decay is a function of time, and less time elapses in the reference frame of the moving muons. In fact, time passes about 22 times more slowly in the moving muons' reference frame. If 2.2 μs have passed according to the scientists' clock, then only 0.1 μs have passed in the moving muon reference frame, which is a time interval much shorter than the half life of the muons. The scientists see that far less than one-half of the muons have decayed in 2.2 μs, and conclude that the moving muons decay more slowly. In the reference frame of the muons, the decay rate is unchanged. Time dilation explains the discrepancy.

You can confirm the "22 times" ratio using the time dilation equation. The proper time is being measured by the muons' "clock", which is observed via their decay rate.

Experiments with highly precise clocks have also confirmed Einstein's conclusions concerning the effect of motion on time. For instance, scientists have measured a difference in the time interval measured by a clock in a plane to that measured by a clock on the ground. Einstein's general theory of relativity and its predictions figure prominently into the results, but his two theories account for the discrepancies between the time intervals measured by the clocks.

Do you experience time dilation? Yes, but very small amounts of it, given how slowly you move compared to the speed of light. Using the equation to the right, you could determine that if you travel for an hour in an airplane flying at 1000 km/h, you would have aged about 0.0000000015 (that is, 1.5×10^{-9}) seconds less than a person who remained stationary on the ground. To provide another sense of the magnitudes, if you moved at a speed of 90 km/hr away from a twin for a 100-year lifespan, you would have lived 11 microseconds less. On the other hand, if you could move at one-tenth the speed of light during a century of his life, you would be six months younger, and if you moved at 90 percent of the speed of light, 270,000,000 meters every second, you would be 44 years old when he was 100.

equation 1



Time dilation

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

t = "stationary" observer time
t_0 = proper time in "moving" frame
v = speed of reference frame
c = speed of light

$$t = \gamma t_0$$
$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

example 1

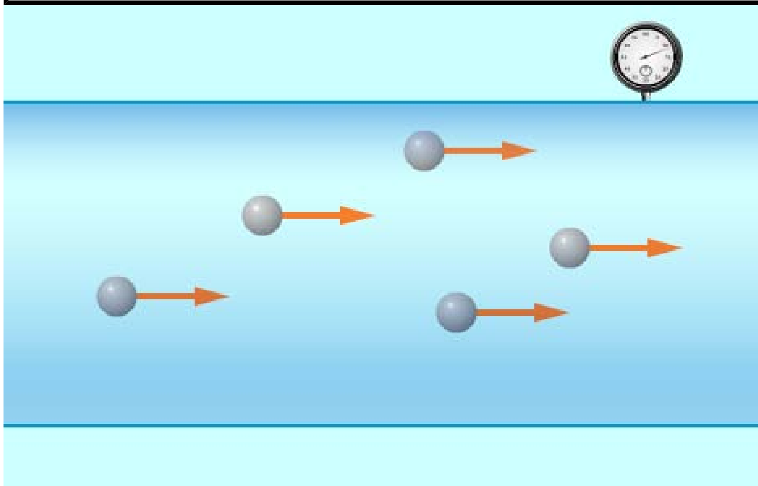
20 seconds have passed between two events on the rocket. How much time has passed on Earth?

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

$$t = \frac{20\ \mathrm{s}}{\sqrt{1 - (0.60c)^2/c^2}} = \frac{20\ \mathrm{s}}{\sqrt{1 - 0.60^2}}$$

$t = 25$ s

40.9 - Interactive checkpoint: time dilation



Certain short-lived elementary particles, called *positive pions* (π^+), have a lifetime of 2.60×10^{-8} seconds, on average, before they decay. This lifetime is measured when they are stationary, that is, in the rest frame of the pions.

A cluster of pions is produced, moving at a speed of 2.95×10^8 m/s relative to the laboratory.

What is the average pion lifetime as measured in the laboratory?

How far do they travel in the laboratory during that time interval?

Answer:

$\Delta t_{lab} =$ ______ s

$d =$ ______ m

40.10 - Exploring and deriving time dilation

concept 1

Light clock
Light flash bounces up and down

concept 2

Professor's reference frame
Light moves strictly up and down

concept 3

Katherine's reference frame

We will use the experiment portrayed on the right to derive the equation for time dilation. Following the steps of the derivation may also help you to understand how Einstein's postulates lead inescapably to the conclusion that time dilation occurs.

The experiment uses a light-based clock mounted on a high-speed skateboard. Ordinary clocks use periodic mechanical or electric processes, such as the oscillation of a pendulum or a timing circuit, to establish a unit of time. A *light clock* uses the amount of time it takes a pulse of light to travel a particular distance. The light clock is convenient to use in this scenario, but any clock would record the same result.

A light pulse is emitted from the base of the clock. The pulse reflects off the top of the clock and returns to the bottom. The clock measures time by using the relationship of time to distance and speed. The elapsed time for the up-and-down journey equals the distance the light pulse travels, divided by the speed of light.

In our experiment, let's consider one "tick" of the clock. The light rises from the bottom of the clock, reflects off the top, and returns to the bottom. The professor, who is also on the skateboard, sees the light pulse moving straight up and down, and he measures the distance it travels as being twice the height of the clock.

Now, let's consider what another observer sees. Katherine is standing on the ground and watches the professor and the clock pass by. She also watches the light pulse as it moves. However, she measures a different value for the distance traveled by the light pulse. She sees it not only moving up and down, but also forward. She measures the light pulse moving through the distance indicated by the two lines labeled s in Concept 3.

If this seems confusing, just think of a friend riding a train, throwing a ball straight up and down. From your friend's perspective, the ball just travels up and down. From a vantage point on the ground outside, you would see the ball moving horizontally at the same time it is moving up and down.

The clock uses light, which according to Einstein's second postulate has a **constant speed** independent of any frame of reference. Einstein's first postulate states that the **laws of physics are the same** in any inertial reference frame. That is, in our scenario, both observers can use the same equation: Time equals the distance traveled by the pulse, divided by the speed of light.

Having explained the experiment, we will now analyze it algebraically, calculating the distance traveled by the clock and the time interval required for one tick of the clock in each frame of reference. This will enable us to derive the equation for time dilation shown in Equation 1.

Variables

We will use the triangle shown to the right in Equation 1 to relate displacement, speed and a time interval. The triangle reflects half of one tick of the light pulse, as it moves from the bottom to the top. The clock moves L horizontally during half of one tick, and Katherine observes the light moving a distance s.

	measured by Katherine	measured by professor
clock's horizontal displacement, half tick	L	0 m
distance light pulse moves, half tick	s	h
clock's speed	v	0 m/s
height of clock	h	h
elapsed time	t	t_0
speed of light	$c = 3.00\times10^8$ m/s	

Strategy

1. Use the Pythagorean theorem to determine the distance the light pulse moves in one tick as measured by Katherine.
2. Use the fact that distance equals the product of speed and time to replace the distances in the diagram.
3. Simplify the equation.

Mathematics principle

We will use the Pythagorean theorem

$$c = \sqrt{a^2 + b^2}$$

Step-by-step derivation

In Katherine's reference frame, the light traces out the hypotenuses of two right triangles. We use the Pythagorean theorem to calculate the hypotenuse. The triangle's height is calculated using the speed of light and the time measured by the professor. Its base is calculated using Katherine's measurement of the clock's speed.

Step	Reason
1. $s = \sqrt{h^2 + L^2}$	Pythagorean theorem
2. $2s = \sqrt{(2h)^2 + (2L)^2}$	light travels two hypotenuses
3. $2h = ct_0$	definition of speed
4. $2L = vt$	definition of speed
5. $2s = \sqrt{(ct_0)^2 + (vt)^2}$	substitute equations 3 and 4 into equation 2

Light moves up, down, horizontally

concept 4

Time measurements differ

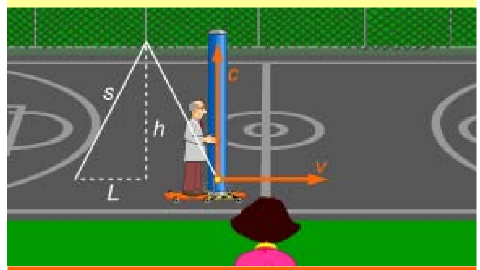
Time = distance / speed of light

Katherine, professor measure different distances

Speed of light is constant

Katherine, professor measure different time

equation 1



Equation for time dilation

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

t = "stationary" observer time

t_0 = proper time in "moving" frame

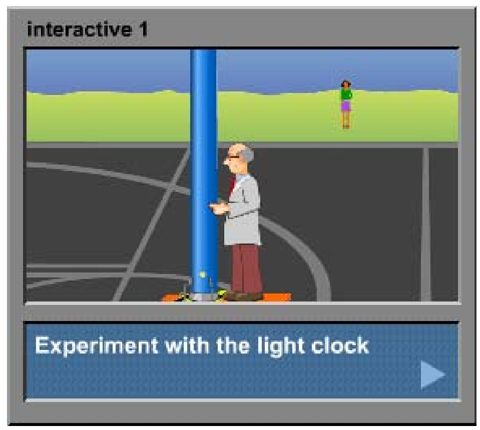
v = speed of reference frame

c = speed of light

We have one last distance left, and again we substitute for it using the speed equation. Then a series of algebraic steps yield the equation for time dilation.

Step	Reason
6. $2s = ct$	definition of speed
7. $c^2t^2 = (ct_0)^2 + (vt)^2$	substitute equation 6 into equation 5; square both sides
8. $c^2t^2 - v^2t^2 = c^2t_0^2$	expand and re-arrange
9. $t^2 = \dfrac{c^2 t_0^2}{(c^2 - v^2)} = \dfrac{t_0^2}{(1 - v^2/c^2)}$	divide both sides by $c^2 - v^2$; simplify
10. $t = \dfrac{t_0}{\sqrt{1 - v^2/c^2}}$	take square root

40.11 - Interactive problem: Experiment with the light clock



In this simulation, you experiment with a light clock. The professor rolls across a basketball court with a light clock on his skateboard, while Katherine watches from the side. You will view the light clock in operation from the professor's reference frame, and from the student's reference frame. You are asked to calculate the time measured by each observer for the professor to cross the basketball court.

The simulation launches in the professor's reference frame, where the light clock is stationary. Press GO and watch the basketball court and the background pass by. A counter will record the number of light clock cycles, the number of round trip journeys made by the light pulse. The clock is 3.0 m tall.

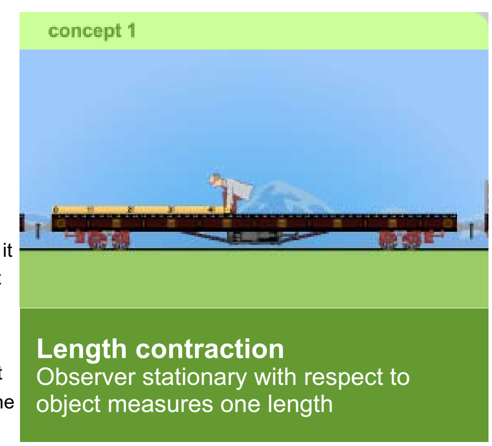
Then press the tab labeled "Student's reference frame" and press GO again. You will see the same series of events from Katherine's reference frame. The simulation displays the path of the light and indicates some key distances in a fashion similar to the derivation of the prior section.

If you asked the professor how long it took him to cross the basketball court, what would he say? What if you asked Katherine the same question? Check your answers by entering them in the simulation.

Now a thought question for you, foreshadowing a future topic. Katherine measures the length of this court as 24.0 m (alas, she does not have an NBA standard court). You can use this length and the time she measures to determine the professor's speed. This is the same speed the professor would measure of the ground moving beneath him.

Let's say the professor decided to use the time he measures and this speed to determine how long the court is. How long does he think the court is? Does he think it is shorter or longer than 24.0 m?

40.12 - Length contraction



Length contraction: The length of an object is less when it is moving relative to an observer than when it is stationary.

Einstein's postulates require that time dilates − observers measure different time intervals between two events when their reference frames are in relative motion. The postulates also require that length "contract". The length of an object will be less when it is moving relative to an observer than when it is stationary. (You will see that the effect is far too small to measure for everyday speeds, which is why you do not notice this effect.)

Consider the train scenario on the right. The train is moving rapidly past an observer at $0.8c$. This is not an everyday speed for even a bullet train. The professor is onboard the train, at rest relative to it. He measures the length of his car as 10 meters. This is the car's *proper length*, the length measured by an observer stationary relative to the length being measured. The use of proper here is analogous to the use of proper in "proper time".

Sara, standing on the ground, measures a different value. She measures the train car as being just six meters long. The relative motion of the two observers causes the differing measurement.

The equation on the right quantifies length contraction. This equation converts the proper length measured by an observer who is stationary relative to the length being measured, to the length observed by an observer who views the object as moving.

As with time dilation, you face two challenges: First, believing it to be true, since it defies your intuition and experience, and second, understanding the notation for the equation. We are not sure how much we can change your intuition, but at the least we can work on the notation. To stress the notation once more: The proper length (L_0) on the right side of the equation is the length measured in a reference frame stationary relative to the object being measured. In the illustration, it is the professor measuring the length of the car.

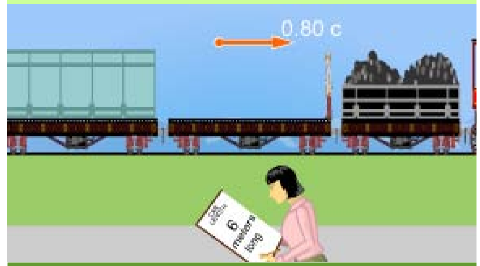
The quantity L on the left-hand side of the equation represents the length that will be measured by a person who views the object (and its reference frame) as moving. This person observes the object as moving at velocity v.

Length contraction occurs only along the direction of the relative motion. In the scenario in this section, only the car's length changes, not its height nor depth. If the train were carrying a light clock, the **vertical** dimension of the clock, used by onboard travelers to measure time intervals, would not be affected by length contraction.

Does the moving object "really" contract? Well, to pose an analogous question from everyday life, do faraway objects really get smaller, as they appear to our eyes? The observations and measurements made by different observers do differ. In discussions of relativity, much focus is placed on observers and measurements, since experiments show effects such as time dilation and length contraction are real.

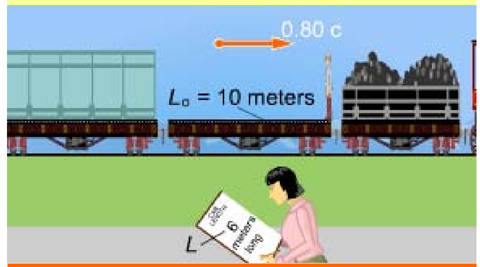
On the other hand, we are taking liberties in showing the object's contraction in these illustrations. If the train car were photographed from a station as it passed by, it would appear rotated. The rear of the car is farther away from the stationary observer, and light from the rear will take longer to reach her than light from the front of the vehicle. The light that arrives simultaneously for the observer is emitted at different instants in time from the car, which moves during that interval of time. Using geometrical arguments, it can be shown that this effect creates the perception of rotation.

concept 2



Length contraction
Observer viewing moving object measures shorter (contracted) length

equation 1

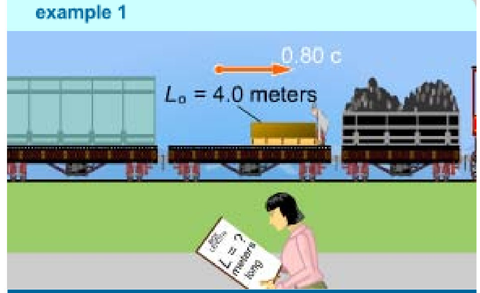


Length contraction

$$L = L_0\sqrt{1 - v^2/c^2}$$

L = length measured by observer watching object pass by
L_0 = proper length (measured by observer stationary relative to object)
v = speed of reference frame
c = speed of light

example 1



How long is the box on the train from Sara's perspective? The professor measures its length as 4.0 meters.

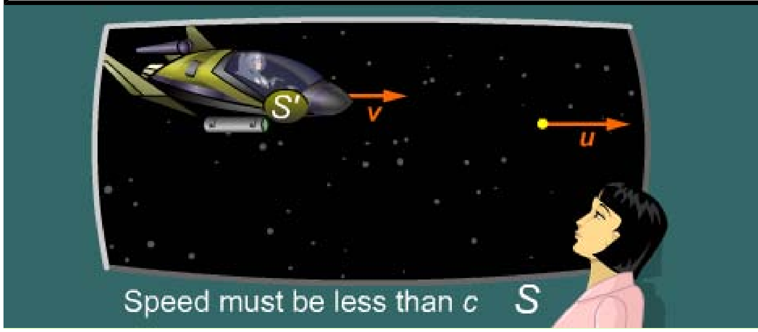
$$L = L_0\sqrt{1 - v^2/c^2}$$

$$L = (4.0\text{ m})\sqrt{1 - (0.80c)^2/c^2}$$

$$L = 4.0\sqrt{1 - (0.80)^2} = 4.0\sqrt{1 - 0.64}$$

$$L = 4.0(0.6) = 2.4\text{ m}$$

40.13 - Relative velocity at relativistic speeds



concept 1

Calculating relativistic velocities
Speed must be less than *c*

Einstein's second postulate states that the speed of light is a constant in all reference frames. It can also be shown that accelerating any massive object to the speed of light requires an infinite amount of energy. The combination of these two considerations would seem to imply that the measured speed of any object in any frame of reference could not reach − or exceed − the speed of light.

However, you can imagine cases where something might seemingly exceed the speed of light. Consider the illustration above, with an astronaut in reference frame S' viewed by an observer in reference frame S. The astronaut fires a flare from her spacecraft at $0.60c$. The observer in S watches the spacecraft pass by at $0.75c$. Can we add these two velocities, and conclude that the flare is moving at $1.35c$?

As you may suspect, the answer is no. If we were ignoring relativistic effects, we could. For instance, if the spacecraft were moving at 60.0 m/s and fired the flare at 75.0 m/s, we could conclude that the earthbound observer would observe the flare moving at 135 m/s, and we would be right, to the level of precision stated in the scenario. If we factored in relativistic effects, they would be *extremely* minor at these speeds.

However, at speeds near the speed of light, or to take full account of relativistic effects even at lower speeds, the two speeds cannot just be added. The Lorentz velocity transformation equation on the right must be used. In this scenario, it means the space station observer measures a velocity of $0.93c$, a value we calculate in the example problem.

The velocity transformation equation shown converts velocity components along the x-axis (or, if you like, the x' axis) from one reference frame to another. It is derived from the prior Lorentz transformation equations and is used under the same conventions and assumptions. Similar-looking expressions, not covered here, transform velocity components in the y and z directions. You may be surprised that there is a difference, since lengths along those axes are the same between frames. However, keep in mind that time is not absolute, so there will be a difference in measured velocities along the directions that are perpendicular to x.

The equation can also be used to confirm that c is a constant in any reference frame. If instead of a flare we consider a beam of light moving at $u' = c$, the equation can be simplified to show that u (the velocity of the light measured on Earth) also equals c.

equation 1

Lorentz velocity transformation equation

$$u = \frac{u' + v}{1 + u'v/c^2}$$

u = velocity of object measured in S
u' = velocity of object measured in S'
v = velocity of S' measured from S
c = speed of light

example 1

In *S′*, the flare is moving at 0.60*c*. How fast is it moving in *S*?

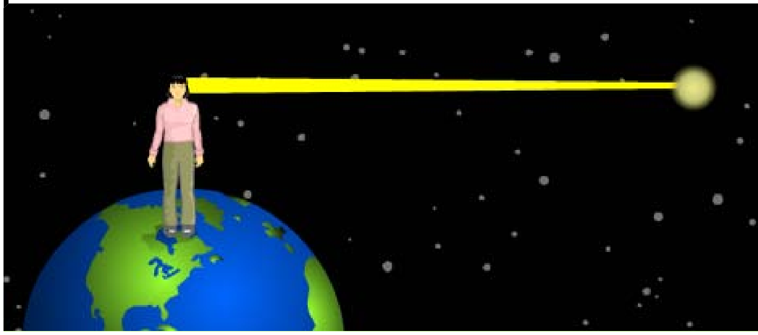
$$u = \frac{u' + v}{1 + u'v/c^2}$$

$$u = \frac{0.60c + 0.75\,c}{1 + (0.60c\,)(0.75c\,)/c^2}$$

$$u = \frac{(0.60 + 0.75)c}{1 + (0.60)(0.75)}$$

$$u = 0.93c$$

40.14 - Doppler shift for light



concept 1

Doppler shift: light
Motion changes light frequency (color)
Shift depends on velocity between source, observer

The Doppler effect is frequently thought of in terms of sound. The siren on a police car heading toward you will sound different than when it is moving away from you. This effect occurs because the motion of the siren changes the frequency of the sound waves that reach you.

The change in the frequency depends on whether you move toward the source of a sound, or the source moves toward you. There is an asymmetry because sound travels through a medium, and it matters which reference frame − yours or the source's − is stationary with respect to the medium.

In contrast, light requires no medium. With light (and other electromagnetic radiation), the Doppler effect depends solely on the relative motion of the source of the light and the observer of the light. In some ways, this makes the Doppler shift "easier" to calculate with light.

On the right, we show equations for calculating the change in perceived frequency due to the Doppler shift. The symbol f_0 represents the frequency of the light as it is emitted from the source (or as it would be perceived by an observer stationary relative to the source). This is called the *proper frequency*. The symbol f represents the frequency perceived by an observer in a reference frame in which the source is moving at speed v. The speed is always positive. The two equations are the same except for the signs; the signs differ for the reasons discussed below.

When the observer and the source are moving toward one other, the frequency of the light seen by the observer will increase. The speed of the wave − the speed of light − must remain constant. Since the wavelength equals the speed of light divided by the frequency, then as the frequency increases, the wavelength decreases. When the source is moving away from the observer, the opposite effect occurs: The wave frequency decreases and the wavelength increases.

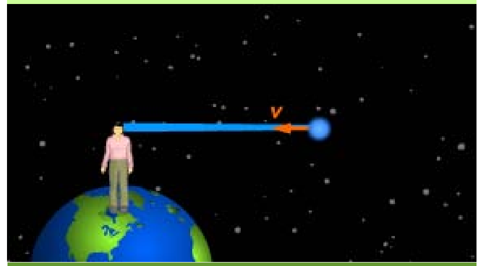
With light, decreasing wavelength due to the approach of the light source is called a *blue shift* since shorter wavelengths occur near the blue end of the visible spectrum. Increasing wavelength due to the recession of the light source is called a *red shift*. Almost all distant stars and galaxies, in whatever direction you look, exhibit red shift. In fact, the farther away they are, the greater the red shift. This is used as evidence that the universe is expanding.

The Doppler effect proved to be crucial to astronomers trying to understand the nature of the universe. Einstein's work provided grounds for believing that the universe was expanding, an implication that proved quite troubling to Einstein himself, who "corrected" his equations to provide for a static universe. But several years after Einstein performed this correction, the American astronomer Edwin Hubble (1889−1953) showed that distant stars and galaxies exhibit a red shift.

Astronomers like Hubble were trained to expect certain patterns in the light emanating from stars. Elements like calcium in the atmospheres of stars absorb certain wavelengths of light; these show up as "gaps" − *spectral lines* − in the light emitted from stars. Hubble expected those spectral lines to occur at certain wavelengths. Instead, he found them "red shifted" to different wavelengths. Although Hubble knew that the Doppler shift could account for these changes, he and his colleagues were astonished that stars in every direction were moving away.

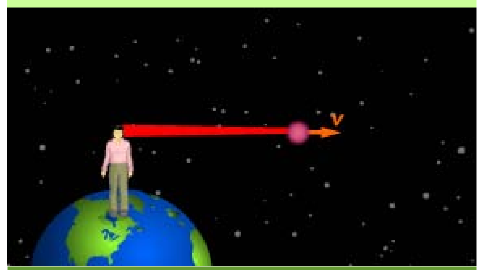
The debate about the nature and fate of the universe continues; but the red shift of starlight provided crucial data to scientists that caused them to further examine the possibility that the universe is expanding. The second sample problem to the right challenges you to compute the speed at which a galaxy is moving away from the Earth based on the shift in its observed wavelength.

concept 2



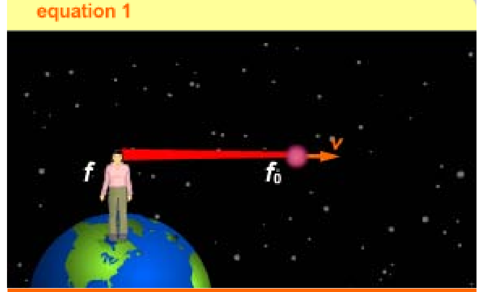
Light source closing in
Blue shift

concept 3



Light source moving away
Red shift

equation 1



Doppler shift

source/observer closing:

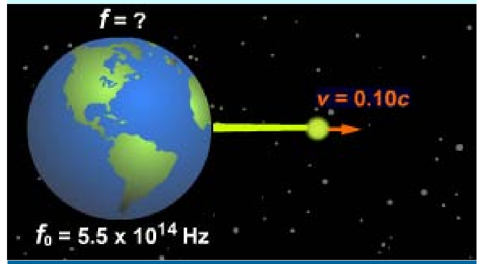
$$f = f_0\sqrt{\frac{1+v/c}{1-v/c}}$$

source/observer separating:

$$f = f_0\sqrt{\frac{1-v/c}{1+v/c}}$$

f = observed frequency
f_0 = proper frequency of source
v = source/observer relative speed

example 1



What frequency of light is observed on Earth?

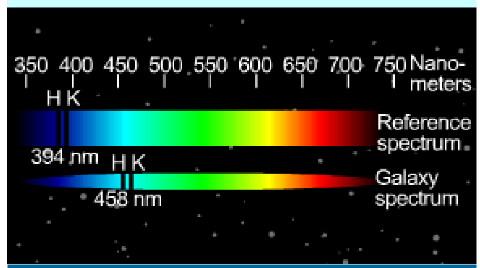
$$f = f_0\sqrt{\frac{1-v/c}{1+v/c}}$$

$$f = (5.5\times10^{14})\sqrt{\frac{1-(0.10c)/c}{1+(0.10c)/c}}$$

$$f = (5.5\times10^{14})\sqrt{\frac{0.9}{1.1}}$$

$f = 5.0\times10^{14}$ Hz

example 2



How fast is the galaxy receding?

Wavelength λ_0 of H K lines is 394 nm

Redshifted wavelength λ is 458 nm

$f = c/\lambda$

$$f = f_0\sqrt{\frac{1-v/c}{1+v/c}}$$

$$\sqrt{\frac{1-v/c}{1+v/c}} = \frac{f}{f_0} = \frac{c/\lambda}{c/\lambda_0} = \frac{\lambda_0}{\lambda}$$

$$\frac{c-v}{c+v} = \frac{\lambda_0^2}{\lambda^2}$$

$$v = \left(\frac{\lambda^2-\lambda_0^2}{\lambda^2+\lambda_0^2}\right)c = \left(\frac{458^2-394^2}{458^2+394^2}\right)c$$

$v = 0.149c = 4.47\times10^7$ m/s

40.15 - Relativistic linear momentum

In discussing relativistic linear momentum, and the equations that describe it, we start with three premises.

First, the same laws of physics hold true in any inertial reference frame. This is Einstein's first postulate. If a property like momentum is conserved in one inertial reference frame, it must be conserved in any other inertial reference frame.

Second, Einstein's second postulate, that the speed of light is constant. These two postulates were used to deduce equations for time dilation, length contraction and so on.

Third, we expect that at much lower speeds, the equations in this section will essentially reduce to the "classical" ones that work perfectly well at low speeds. For instance, the momentum of an object moving at say 30 m/s should very, very closely equal mv. This has held true with prior equations in this chapter. At speeds much less than the speed of light, the equations predict results that accord with "classic" mechanics equations.

These three premises are satisfied by the relativistic equation for momentum at the right. This extended equation for momentum is required because it can be shown that if Newtonian, or classical, momentum is conserved in one reference frame, then when the velocities are converted to the values that would be observed in another frame, the total momentum is never conserved in that second reference frame.

Specifically, if a collision occurs in a moving reference frame S', and the velocities are converted to the values that would be observed in S, the quantity mv before and after the collision is changed by the collision. Using the equation shown to the right ensures this will not occur, and that momentum is conserved when the collision is observed from any reference frame.

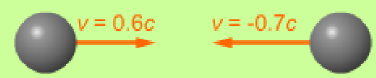
The larger the relative speed of the frames, the greater the discrepancy between the classical momentum and the relativistic momentum will be. It turns out that the Lorentz factor γ appearing in the definition of momentum shown in Equation 1 provides exactly the right corrective factor for momentum to be conserved in both S and S', no matter how fast the frames move with respect to each other. (The variable u in the formula refers to the speed of an object in its frame, not the speed of one frame with respect to another.)

The extended momentum equation obeys the first two premises stated above. It also obeys the third. At velocities much less than the speed of light, momentum approaches mv, since dividing the square of a velocity like 300 m/s by the square of speed of light means the relativistic effect is near nil.

As the example on the right shows, at half the speed of light, relativistic effects increase momentum by 15% over its classical value. The effect increases significantly as an object moves at speeds closer to the speed of light: At 99% of the speed of light, momentum is about seven times its classical value.

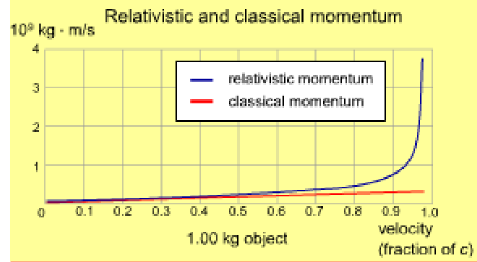
This phenomenon is of great importance in the modern physics research done with large particle accelerators, such as the Fermilab accelerator in Illinois, or the CERN accelerator in France and Switzerland. Objects move near the speed of light in these accelerators, and calculating their momentum requires the use of the equation shown in this section.

concept 1



Relativistic momentum
New equation for momentum required for high velocities

equation 1

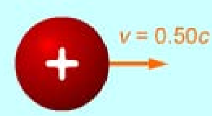


Equation for relativistic momentum

$$p = \frac{mu}{\sqrt{1 - \frac{u^2}{c^2}}} = \gamma mu$$

p = momentum
m = mass
u = velocity of object
c = speed of light
γ = Lorentz factor

example 1



What is the momentum of the proton? Its mass is 1.67×10⁻²⁷ kg.

$$p = \frac{mu}{\sqrt{1 - \frac{u^2}{c^2}}}$$

$$p = \frac{(1.67 \times 10^{-27}\ \text{kg})(1.50 \times 10^{8}\ \text{m/s})}{\sqrt{1 - ((0.500c)^2/c^2}}$$

$p = 2.89\times10^{-19}$ kg·m/s

(Relativistic effects add 15%)

40.16 - Mass and energy

The world's "most famous equation" is shown in Equation 1: *Rest energy* (also known as *mass energy*) equals mass times the speed of light squared. By "rest energy", we mean the energy equivalent of its mass.

If a mass is moving, its total relativistic energy can be calculated as the sum of this value and its kinetic energy. The equation for calculating this total energy is shown in Equation 2. The equation factors in relativistic effects when calculating KE.

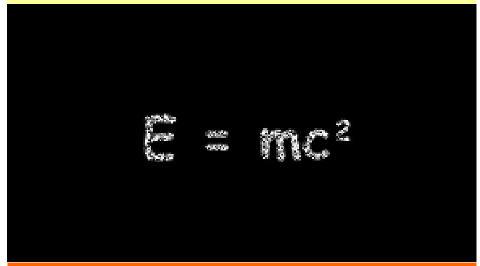
It is interesting to consider Equation 1 in the context of the Sun. The Sun becomes less massive as its mass is transformed by fusion reactions into the energy it radiates. It radiates 3.91×10^{26} joules of energy per second, and as a consequence, it "loses" more than four billion kilograms per second! (But do not worry! Its total mass is 1.99×10^{30} kg).

Earlier scientists had proposed the conservation of mass in chemical reactions. This textbook has discussed the conservation of energy. At non-relativistic speeds, these principles hold true to a high degree of accuracy. Einstein showed both mass and energy must be considered when applying conservation principles.

Thermonuclear processes inside the sun convert a very large quantity of mass into a stupendous quantity of energy every second. Much more comprehensible amounts of matter yield tremendous amounts of energy as well. For example, if one gram of matter were converted entirely to energy, it would produce the energy equivalent of combusting fifteen thousand barrels of oil. On the dark side of things, converting into energy a tiny fraction of the 2 kg of matter in a nuclear bomb can produce a blast equivalent to the chemical energy released from the explosion of 50 billion kilograms of dynamite.

Atomic particles are often subject to relativistic effects because of their high speeds. For convenience, the energy and mass of such particles is often measured in units other than joules and kilograms. An electron volt (eV) equals the change in the potential energy of an elementary charge (1.60×10^{-19} C) when it moves through a potential difference of one volt. Electron volts are convenient measures of particle energy. Particle mass is often measured in units of eV/c^2, or energy divided by the speed of light squared. The fact that these are mass units follows from the equation $E = mc^2$.

equation 1

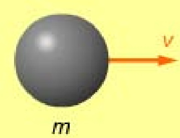


Energy at rest

$$E_0 = mc^2$$

E_0 = rest energy
m = mass
c = speed of light

equation 2

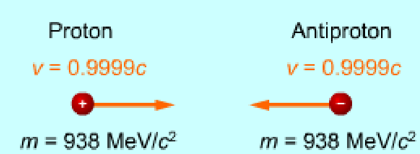


Total energy

$$E = \gamma mc^2$$

E = total energy
γ = Lorentz factor
m = rest mass
c = speed of light

example 1



When a proton and antiproton collide, they annihilate each other and their mass energy is converted into a pair of gamma

ray photons. Calculate the total energy of the gamma rays.

$$\gamma = \frac{1}{\sqrt{1 - u^2/c^2}}$$

$$\gamma = \frac{1}{\sqrt{1 - (0.9999c)^2/c^2}} = 70.7$$

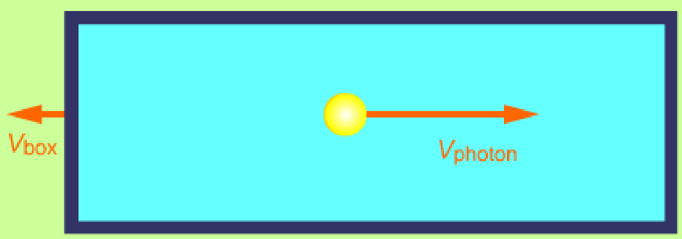
$E = \gamma mc^2$

$E = (70.7)(938\ \text{MeV}/c^2)c^2$

$E = 6.63\times10^4$ MeV

$2E = 1.33\times10^5$ MeV

40.17 - "E equals mc squared": a thought experiment



concept 1

$E = mc^2$: a thought experiment
Photon emitted and absorbed in box
Center of mass stays same
· Energy equivalent to mass

In a brilliant thought experiment, Einstein demonstrated how energy and mass are related. The experiment is illustrated in the animation above.

To explain Einstein's thought experiment, we first note that in 1864, Maxwell's publication on electromagnetic radiation showed that "electromagnetic waves" carry momentum. Maxwell's conclusion was surprising to his peers, since electromagnetic radiation (like light) was thought of as solely energy. However, his work showed it was displaying a characteristic of mass: It had momentum. (In a space of fifty years, scientists had to move from light being considered as a wave moving through a medium, to a particle capable of moving through a vacuum at a constant speed.) Other research, including work by Einstein that won him a Nobel prize, caused scientists to coin the term *photon* to refer to a light "particle".

Einstein reasoned that the energy of a photon must be equivalent to a certain amount of mass, which could be related to its momentum. He formulated the *gedanken* experiment described in this section to explain his ideas, and this experiment can be used to demonstrate a mathematical relationship between energy and mass.

First, imagine a stationary box floating in deep space. A photon is emitted from inside the left side of the box and it moves toward the right. For momentum to be conserved the box must recoil as the photon is emitted, and move slowly to the left. The photon then collides inelastically with the other side of the box (the photon is absorbed), transferring its momentum to the box, which causes the box to stop moving. The total momentum of the system before the photon is emitted, and after it is absorbed, equals zero.

Since no external force acts on this system, the center of mass must stay in the same location. However, the box has moved. How can the movement of the box be reconciled with the center of mass of the system remaining fixed?

Einstein resolved this apparent contradiction by proposing that there must be a **mass equivalent** to the energy of the photon. By considering the mass equivalent of the photon as having moved from the left to the right side of the box, the center of mass stays constant. It is like system consisting of a person walking on a floating raft. As the person (photon) moves to the right, the raft (box) moves to the left, but the center of mass stays in the same location.

In the following derivation we perform the algebraic steps corresponding to each stage of Einstein's thought experiment, and so arrive at $E = mc^2$.

Variables

The original position of the box's center of mass is given in a coordinate system whose origin coincides with the left side of the box at the beginning of the experiment.

momentum of photon	p
energy of photon	E
speed of light (and of the photon)	c
mass equivalent of photon	m
mass of box	M
length of box	L
original position of box's center of mass	$\bar{x}_0$
speed of box during photon transit	v
distance moved by box during transit	Δx
time required for photon transit	Δt

Strategy

1. State an equation expressing the conservation of momentum in the system as the box recoils from the emission of the photon.
2. Apply the definition of speed to derive an expression for the change $M\Delta x$. If we ignored the mass equivalent of the photon, this would be the change in the position of center of mass of the system.
3. Posit a mass m equivalent to the energy of the photon so that, in moving to the right, it cancels the leftward shift in the center of mass due to the motion of the box. Use the constancy of the system's overall center of mass to determine a quantitative relationship between the mass m, the photon's energy E, and c^2.

Physics principles and equations

For the momentum of our photon, we will use Maxwell's expression for the momentum of an electromagnetic wave having a given energy. The momentum is

$$p = \frac{E}{c}$$

The box will not move very rapidly, so we can use the classical definition of momentum

$$p = mv$$

The definition of speed

$$v = \frac{\Delta x}{\Delta t}$$

We use a formula for calculating the x coordinate of the center of mass of a system of two objects

$$\bar{x} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}$$

Step-by-step derivation

We state an equation reflecting the conservation of momentum as the box recoils from the emission of the photon at the beginning of the thought experiment.

Step	Reason
1. $p = \frac{E}{c}$	momentum of wave/photon
2. $p = Mv$	momentum of box
3. $Mv = \frac{E}{c}$	conservation of momentum

We now find the change $M\Delta x$ in the contribution of the box to the system's center of mass.

Step	Reason
4. $v = \dfrac{\Delta x}{\Delta t}$	definition of speed
5. $M\dfrac{\Delta x}{\Delta t} = \dfrac{E}{c}$	substitute equation 4 into equation 3
6. $M\Delta x = \dfrac{E}{c}\Delta t$	rearrange
7. $\Delta t = \dfrac{L}{c}$	definition of speed
8. $M\Delta x = \dfrac{EL}{c^2}$	substitute equation 7 into equation 6

Since the overall center of mass of the system does not change, we can find a relationship between the changes to the center of mass due to the displacements of the box and the photon. This leads to Einstein's equation $E = mc^2$.

Step	Reason
9. $\dfrac{M\bar{x}_0 + m\cdot 0}{M+m} = \dfrac{M(\bar{x}_0 - \Delta x) + mL}{M+m}$	center of mass does not change
10. $M\bar{x}_0 = M\bar{x}_0 - M\Delta x + mL$	simplify
11. $mL = M\Delta x$	solve for mL
12. $mL = \dfrac{EL}{c^2}$	substitute equation 8 into equation 11
13. $E = mc^2$	solve for E

Einstein actually stated his famous equation in this form: "if a body gives off the energy E in the form of radiation, its mass diminishes by E/c^2." Note that a photon has no mass; its mass **equivalent** is E/c^2. The mass of an object like the box is diminished by the release of energy.

A clever student will note that since the box is moving, the distance traveled by the photon is slightly less than the box's length. A challenging physics professor could ask the student to conduct this thought experiment even more carefully, including other factors like the reduced mass of the box during the photon's transit. The result of conducting the thought experiment with more rigorous precision is the same.

40.18 - Total energy: relativistic vs. Newtonian kinetic energy

Einstein's formula for the energy equivalent to a mass at rest is $E_0 = mc^2$. For a mass in motion his expression for the *total energy* is $E = \gamma mc^2$. The difference is due to the object's energy of motion, its kinetic energy. The "classic" definition for KE is $\frac{1}{2}mv^2$, but below, we determine its KE_{rel}, its kinetic energy factoring in relativistic effects. We do so by subtracting its energy at rest from its total energy.

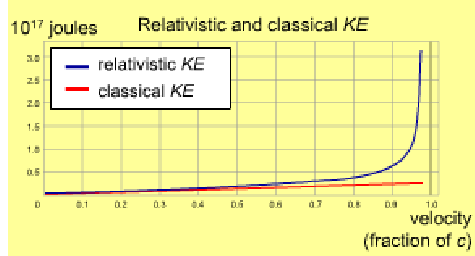
$$KE_{\text{rel}} = E - E_0 = \gamma mc^2 - mc^2 = (\gamma - 1)mc^2$$

In the previous section we followed Einstein's thought experiment to derive the formula $E_0 = mc^2$. We can now justify the formula for the total energy E if we can justify the equation for KE_{rel}. A full derivation of this equation is beyond the scope of this book, but we will show that it reduces to the Newtonian formula for kinetic energy when the speed u of an object in a reference frame is small compared to the speed of light. We will call Newtonian kinetic energy KE_{class} , for the "classic" definition of kinetic energy.

Variables

Lorentz factor	γ
speed of object in reference frame	u
speed of light	c
Newtonian kinetic energy	KE_{class}
relativistic kinetic energy	KE_{rel}
mass of object	m

equation 1



Relativistic kinetic energy

$$KE_{\text{rel}} = (\gamma - 1)mc^2$$

$KE_{\text{rel}} = KE$, factoring in relativistic effects
m = mass
c = speed of light
u = velocity

Strategy

1. Expand the Lorentz factor γ using the binomial expansion formula. Use this to approximate $\gamma - 1$ for nonrelativistic values of u.
2. Substitute the nonrelativistic approximation of $\gamma - 1$ into the formula for relativistic KE, and observe that it reduces to Newtonian KE.

Physics principles and equations

We use the definition of the relativistic kinetic energy as stated on the right.

The definition of Newtonian kinetic energy

$KE = ½mu^2$

Mathematics principle

In addition we will use the following form of the binomial expansion, valid for any power n and for "small" values of ε ($|\varepsilon| < 1$),

$$(1-\varepsilon)^n = 1 - n\varepsilon + \frac{n(n-1)}{2!}\varepsilon^2 - \frac{n(n-1)(n-2)}{3!}\varepsilon^3 + \ldots$$

Step-by-step derivation

We use the binomial expansion to get an approximation to the Lorentz factor γ.

Step	Reason
1. $\gamma = \dfrac{1}{\sqrt{1-u^2/c^2}}$	definition of Lorentz factor
2. $\gamma = (1-u^2/c^2)^{-1/2}$	rewrite using exponent
3. $\gamma = 1 - (-\frac{1}{2})\frac{u^2}{c^2} + \frac{3}{8}\left(\frac{u^2}{c^2}\right)^2 + \ldots$	binomial expansion
4. $\gamma \approx 1 - \frac{1}{2}\frac{u^2}{c^2}$	u^2/c^2 is very small
5. $\gamma - 1 \approx \frac{1}{2}\frac{u^2}{c^2}$	subtract 1 from both sides

In the following steps, we substitute the approximation to $\gamma - 1$ into the formula for relativistic KE, and simplify the result.

Step	Reason
6. $KE_{rel} = (\gamma - 1)mc^2$	relativistic kinetic energy
7. $KE_{rel} \approx (\frac{1}{2}\frac{u^2}{c^2})mc^2$	substitute approximation 5 into equation 6
8. $KE_{rel} \approx ½mu^2$	simplify
9. $KE_{class} = ½mu^2$	definition of kinetic energy
10. $KE_{rel} \approx KE_{class}$	substitute equation 8 into approximation 7

In the graph in Equation 1 on the right above, we show the difference between using the Newtonian equation for kinetic energy and using an equation that includes relativistic effects. The object whose energy is being graphed has a mass of 1.0 kg. The equation for KE_{rel} shows the reason for this difference. As the speed of an object approaches the speed of light, the denominator of the Lorentz factor will approach zero, meaning the value of the factor itself will approach infinity.

The example shows that to accelerate a 10-kilogram mass to $0.99c$ requires a huge amount of energy: 5.5×10^{18} J. This is about a dozen times larger than the energy that would be calculated if the object's KE is calculated using the classical equation, $½mv^2$. Ten thermonuclear bombs would be one way to obtain this much energy.

γ = Lorentz factor $= \dfrac{1}{\sqrt{1-u^2/c^2}}$

example 1

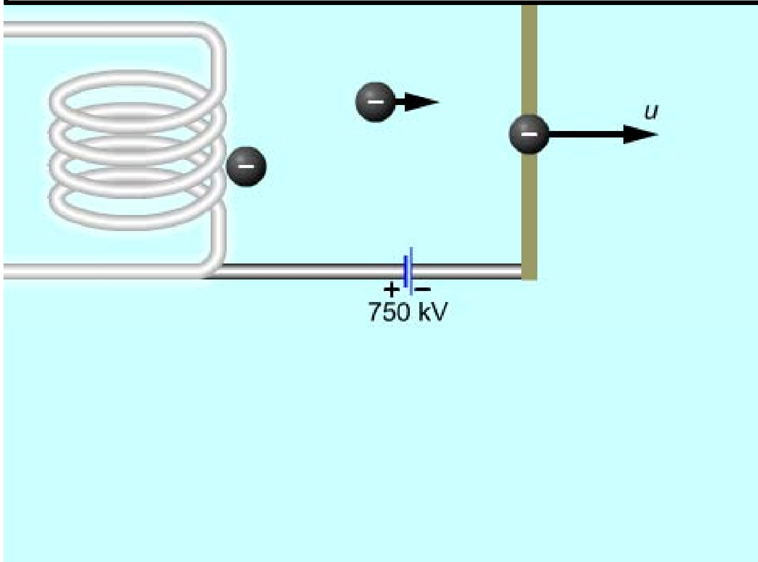
The monkey wants to travel at 0.99*c*. How much energy is required?

$$KE_{rel} = \left(\frac{1}{\sqrt{1-u^2/c^2}} - 1\right)(mc^2)$$

$$KE_{rel} = \left(\frac{1}{\sqrt{1-(0.99c)^2/c^2}} - 1\right)(mc^2)$$

$KE_{rel} = (6.1)(10\text{ kg})(3.0\times10^8\text{m/s})^2$
$KE_{rel} = 5.5\times10^{18}$ J

40.19 - Interactive checkpoint: accelerated electron



The electron has a rest energy of 511 keV. Terence Tightwad, who understands classical physics but has never studied special relativity, proposes a particle accelerator where an electron is accelerated from rest through a potential difference of 750 kV, giving it a kinetic energy of 750 keV.

If the classical expression for kinetic energy always held true, what would be the final speed, *u*, of the electron in terms of *c*?

What is the speed of this electron as predicted by special relativity? (This is what a disappointed Terence would measure, using techniques such as time-of-flight detection.)

Answer:

$u_{\text{classical}} =$ ______ c

$u_{\text{relativity, measured}} =$ ______ c

40.20 - Gotchas

If you measure the speed of light from a star that is moving toward you, you will get a higher value than if you measure the speed of light from a star that is moving away from you. No! You will measure the same speed. The speed of light does not change based on motion of source or observer. This is one of Einstein’s postulates.

You are flying from Earth to Alpha Centauri at a constant speed of 150,000,000 m/s. Given your incredible speed, you should expect some of the laws of physics to change. No, Einstein’s other postulate is that the laws of physics are the same in any inertial reference frame. You are moving at a constant speed, so you are in an inertial reference frame.

I’m watching an airplane flying overhead at 400 km/hr. The pilot throws and catches a ball, measuring the time with a stopwatch. Her stopwatch measures the proper time. Yes, that is correct.

Since you are standing still as you watch the plane pass by, you measure its proper length. No. You have to be at rest relative to an object to measure its proper length. However, in the case of an airplane, the difference is negligible − not so for relativistic speeds!

Einstein was very smart. Yes.

40.21 - Summary

Einstein radically changed our perception of space and time with his two postulates of special relativity. The first postulate states that the laws of physics are the same for observers in any inertial reference frame. The second postulate states that the speed of light in a vacuum is the same in all inertial reference frames.

An event is specified by giving its space and time coordinates. Observers who are in different inertial reference frames will assign different coordinates to the same event.

Observers in a reference frame where two events occur at the same place measure a time interval between those events called the proper time. Observers moving relative to that frame will always measure a longer time interval. This effect is known as time dilation.

The Lorentz transformation equations are used to relate these sets of coordinates. Special relativity predicts that different observers may not agree on the time interval between two events.

The inability to agree on time intervals led to Einstein's thought experiment about the concept of simultaneity. In this thought experiment, he showed that observers in reference frames moving relative to one another would **not** agree that there was zero time interval between separated events − in short, they would disagree about whether events occurred simultaneously or not.

Special relativity also correctly predicts that different observers may not agree on the spatial interval between two events. The length of an object at rest in an inertial reference name is known as its proper length. Observers who are moving relative to that frame will always measure a shorter length than the proper length. This effect is known as length contraction.

The Doppler effect applies to light. The astronomer Hubble used this effect to argue that the universe is expanding.

The effect has a purely relativistic aspect, the transverse Doppler effect, due to time dilation.

Measurements of any object's velocity will vary depending on the inertial reference frame from which the velocity is measured. The Lorentz velocity transformation equation relates velocity values along a line, reported from different frames that move relative to each other along that line.

The intertwining of space and time means that a relativistic definition of linear momentum must be accepted if consistent relations among them are to be retained at both low and high speeds, and for observers in all inertial reference frames.

Special relativity reveals a previously unseen connection between two fundamental concepts: mass and energy. A mass has a tremendous amount of rest energy locked up, ready to be released either in a terrifying chain reaction or in a controlled, useful manner. Meanwhile, the converse is also true: Energy has a mass equivalent and may be transformed into matter.

Special relativity also forces a new perspective on the kinetic energy of an object.

Equations

Time dilation

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

$$t = \gamma t_0$$

Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

Length contraction

$$L = L_0\sqrt{1 - v^2/c^2}$$

Doppler shift

Source/observer closing:

$$f = f_0\sqrt{\frac{1 + v/c}{1 - v/c}}$$

Source/observer separating:

$$f = f_0\sqrt{\frac{1 - v/c}{1 + v/c}}$$

Relativistic momentum

$$p = \frac{mu}{\sqrt{1 - \frac{u^2}{c^2}}} = \gamma mu$$

Rest energy

$$E_0 = mc^2$$

Total energy

$$E = \gamma mc^2$$

Relativistic kinetic energy

$$KE_{\text{rel}} = (\gamma - 1)mc^2$$

Chapter 40 Problems

Chapter Assumptions

Assume the speed of light in a vacuum is 3.00×10^8 m/s for all observers.

Conceptual Problems

C.1 Would a windowless spaceship, rotating at constant speed in deep space, qualify as an inertial reference frame?

○ Yes ○ No

C.2 Jana and Shashi are engaged in a contest to see who can first hit Pluto with a laser beam. At precisely 12 noon, Jana fires her laser from the Earth. In an attempt to beat her, Shashi races by the Earth at $0.50c$ (as measured by Jana), toward Pluto, and fires his beam just as he passes her. Who will win the contest? Ignore any effects of the Earth's atmosphere.

○ Jana ○ Shashi ○ It's a tie!

C.3 Spaceship A and spaceship B are racing toward an airless moon at speeds of $0.15c$ and $0.85c$, respectively. They each fire a pulsed laser as a signaling device, and the speed of each of the pulses is measured on the moon. (a) What is the speed of the pulse received from spaceship A? (b) What is the speed of the pulse received from spaceship B?

(a) ○ $0.15c$ ○ $0.85c$ ○ $1.00c$
(b) ○ $0.15c$ ○ $0.85c$ ○ $1.00c$

C.4 An observer measures two events that occur at the same place and at the same time. For example, two beams of light from two different directions strike a point. (a) Will the two events be simultaneous for all other observers? (b) Explain your answer.

(a) ○ Yes ○ No
(b)

C.5 A scheme for a modern-day Fountain of Youth is introduced by a marketing company. Their idea is to launch their clients on a round trip at relativistic speeds. Their ads proclaim: "After 10 years have elapsed on Earth, you will only have aged 5 years!" Will the passengers enjoy a longer life according to their own biological clocks?

○ Yes ○ No

C.6 A spaceship races by the Earth at a constant velocity of $0.999c$. An astronaut in the ship holds a mirror up and looks at her image in it. (a) Does she see any unusual effects? (b) Describe the effects or explain why she doesn't see any.

(a) ○ Yes ○ No
(b)

C.7 Consider a box on a relativistic train car, moving at $0.866c$ along the positive x axis as viewed from the Earth. The length of the box along the x axis, as measured by Earth observers, will be only half of its proper length. If V is the volume of the box measured by an observer on the train, what is volume of the box as measured by an observer on Earth?

○ V ○ $0.866V$ ○ $V/2$ ○ $V/8$ ○ V/π

C.8 An astronomer on Earth measures the speed of light coming from the star Sirius, and measures $u = c$. Meanwhile, the astronomer also sees spaceship A racing by at $0.715c$ to the right. An observer in ship A sees a scientific research vessel, B, that is moving at $0.650c$ to the left, as measured from reference frame A. The scientists on research vessel B also measure the speed of light coming from the star Sirius. What speed do they measure? Express your answer as a multiple of c. Hint: you should not have to do any calculations.

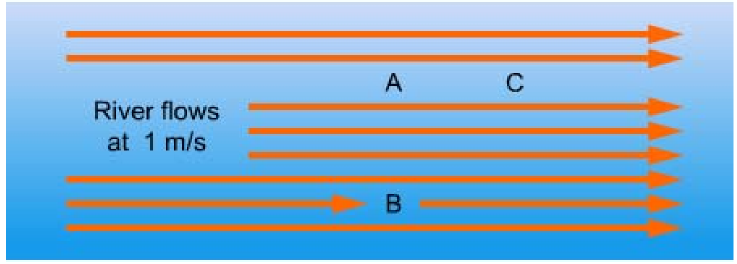
__________ c

C.9 If the speed of light in our universe were only 30.0 m/s, how would this affect life as we know it? Describe at least three consequences.

Section Problems

Section 3 - Light can travel through a vacuum

3.1 Here is an analogy to the Michelson-Morley experiment. Suppose a swimmer can swim at 3.0 m/s in still water (that is the "medium" of travel for the swimmer). Consider three fixed rocks A, B, and C in a flowing river, as shown, such that distance AC is 12 meters and AB is also 12 meters. The river flows at 1.0 m/s. (a) How much time does the swimmer take for round trip ACA, parallel to the river flow? (b) How much time for round trip ABA, perpendicular to the flow?



Explanation of the analogy: By rotating their apparatus, and swapping the role of B and C, Michelson and Morley hoped to measure the speed of the earth with respect to the ether. (The analogy is to try and measure the water speed by making round trip time measurements.) What they found was that there was no difference in the times, no matter what time of year they did their experiments, and no matter what the orientation of the apparatus.

(a) __________ s
(b) __________ s

Section 7 - Interactive problem: Conduct Einstein's simultaneity experiment

7.1 Use the simulations in both interactive problems in this section to answer the following questions. In all cases, use the default setting for the time interval between lightning strikes, 0 ns. (a) Does Katherine observe the lightning bolts as striking the lightning rods simultaneously when the train is stationary? (b) Does the professor? (c) Does Katherine observe the lightning bolts as striking the lightning rods simultaneously when the train is moving? (d) Does the professor?

(a) ○ Yes ○ No
(b) ○ Yes ○ No
(c) ○ Yes ○ No
(d) ○ Yes ○ No

Section 8 - Time dilation

8.1 A rocket ship is 45.0 m long according to measurements made in its rest frame, and moves at 0.385*c* with respect to the Earth. Astronaut Naomi sends a laser pulse from the tail of the rocket to the nose, where it reflects off a mirror, and returns to Naomi. (a) How long does the round trip take, according to Naomi? (b) How long does the round trip take according to Tariq, an observer on Earth?

(a) __________ s
(b) __________ s

8.2 An interstellar spacecraft races by the Earth at a velocity of 1.85×10^8 m/s, as measured by observers on the Earth. The spacecraft's galley cook places a pot of Minute® rice on the stove for exactly 60.0 s, according to his watch. How long will the pot have been on the stove according to the earthbound observers?

__________ s

8.3 An interstellar cruise ship races by a Martian spaceport. According to galactic regulations, the safety strobe lights on the ship's bridge are supposed to blink every 1.50 seconds. A rookie policemartian, based on the planet, issues a ticket, claiming he measured the time interval between blinks as 1.75 seconds. You are called upon by the indignant captain to explain the discrepancy, who maintains that her lights are fully compliant with code. What spacecraft speed will show both the captain and policemartian to be correct?

__________ m/s

8.4 You have probably never noticed time dilation effects in everyday life. Do some calculations to see why this is. (a) Calculate the speed that a friend would have to move with respect to you so that your measured time intervals are 1.00% larger than hers. That is, for each 100 elapsed seconds on her wristwatch, 101 seconds elapse on yours. (b) How many times larger is the speed you just calculated than the speed of an orbiting space shuttle, 8.00×10^3 m/s?

(a) __________ m/s
(b) __________ times the speed of the shuttle

8.5 A pi meson, or pion, is an elementary particle that exists for a brief time. Charged pions at rest have an average lifetime of 26.0 ns. A group of freshly created pions travel, each at a constant speed, an average distance of 28.0 m in the laboratory before decaying. What was the average pion speed as measured in the lab?

_________ m/s

8.6 It is the year 3050, and astronauts are flying to Sirius A, the brightest star in the night sky, which is located 8.6 light-years from Earth (that is, light takes 8.6 years to arrive from Sirius A). (a) Explain how it is possible to make the trip so that the astronauts will age only 7.6 years during the flight. (b) Calculate the constant velocity at which they must travel to do this as measured by Earth/Sirius observers. Express your answer as a decimal fraction of *c*. (c) How long does the trip take according to Mission Control, on Earth?

(a)
(b) _________ c
(c) _________ years

Section 10 - Exploring and deriving time dilation

10.1 Consider a very tall light clock that is 555 meters high. It is mounted vertically on a spacecraft that is moving along the positive *x* axis at 2.00×10^8 m/s, as measured from the Earth. During a certain time interval, an earthbound observer measures the clock moving a distance of 497 meters along the *x* axis, while the light pulse has made a trip from the bottom of the clock to the top. (a) How far has the light pulse traveled, according to an astronaut on the spacecraft? (b) How far has the light pulse traveled, according to earthbound observers? (Your answers to parts a and b should differ.) (c) What is the time interval required for the light pulse to travel from the bottom to the top of the clock, according to the astronaut? (d) What is the time interval required for the light pulse to travel from the bottom to the top of the clock, according to earthbound observers? (e) Calculate the Lorentz factor for the spacecraft. (f) Use the Lorentz factor to calculate the dilated time interval corresponding to the astronaut's measurement in part c. (Your answers to parts d and f should be the same.)

(a) _________ m
(b) _________ m
(c) _________ s
(d) _________ s
(e) _________
(f) _________ s

Section 11 - Interactive problem: Experiment with the light clock

11.1 Use the simulation in the interactive problem in this section to answer the following questions. (a) How long does the professor think it took him to cross the basketball court? (b) How long does Kathering think it took the professor to cross the basketball court? (c) Suppose Katherine measures the length of the court as 24.0 m, and then uses this to calculate the professor's speed. If the professor uses that value for his speed and the time **he** measured, is the professor's measurement of the basketball court's length going to be longer, shorter or the same as Katherine's measurement?

(a) _________ s
(b) _________ s
(c) i. Longer
ii. Shorter
iii. The same length

Section 12 - Length contraction

12.1 A newly-constructed spaceliner boasts a 105 m long and 65.0 m wide soccer field for the enjoyment of its international crew during extended missions. The long dimension of the soccer field is aligned along the ship's direction of flight. As the ship cruises by Earth, FIFA (Fédération Internationale de Football Association) officials on Earth measure the length of the field as 99.0 m and declare the field ineligible since it is shorter than the required 100 m. The captain protests that the discrepancy is due to length contraction. (a) How fast was the ship traveling according to the Earth-bound observers (as a decimal fraction of *c*)? (b) What is the width of the field according to the FIFA officials?

(a) _________ c
(b) _________ m

12.2 You want to visit your friend Ford Prefect, who lives in the vicinity of Betelgeuse, 425 light-years from Earth (that is, light takes 425 years to arrive from Betelgeuse). You hitch a ride on a passing starcruiser, and as measured by Earth observers, you travel so fast that the trip only takes 426 years. (a) How long was the trip according to you? (b) What is the distance between Betelgeuse and Earth, according to you?

(a) _________ years
(b) _________ light-years

Section 13 - Relative velocity at relativistic speeds

13.1 Two spacecraft fly away from Earth, traveling in exactly opposite directions. As measured by an observer on Earth, spacecraft A travels at a speed of 0.392*c* and spacecraft B travels at a speed of 0.885*c*. What is the speed of spacecraft B as measured by A? Express answer as a decimal fraction of *c*.

_________ c

13.2 From Earth's reference frame, you watch a starcruiser fly directly away from the Earth in the positive *x* direction, at a velocity of 0.60*c*. The starcruiser fires two muon torpedoes, one forward and one backward, each at a speed of 0.50*c* as measured by the pilot. (a) What is the velocity of the forward-aimed torpedo, according to you? (b) What is the velocity of the other torpedo, according to you?

(a) _________ c , in the i. positive ii. negative x direction

(b) _________ c , in the i. positive ii. negative x direction

13.3 A ship carrying the Andorian ambassador approaches Earth at 0.55*c*, and a small shuttle is sent from the Earth to greet the diplomat. As seen from the Andorian's ship, the shuttle is approaching at 0.75*c*. What speed do Earth observers measure for the shuttle? Express your answer as a decimal fraction of *c*.

_________ c

13.4 A spaceship passes Earth at a velocity of 0.610*c*. Its length as measured by Earth observers is 157 m. A point-sized ambulance, moving at 0.880*c* according to Earth observers, is speeding in the same direction, and is just about to overtake the spaceship. (a) What is the proper length of the spaceship? (b) How long does it take for the ambulance to pass the ship, according to the ship passengers?

(a) _________ m

(b) _________ s

Section 14 - Doppler shift for light

14.1 Takashi is riding an intergalactic Japanese bullet spacecraft that is heading away from Earth at 0.450*c*, as measured from Earth. An Earth-bound bullet spacecraft is coming toward it, carrying Yuki at 0.213*c*, as measured from Earth. Takashi and Yuki transmit radio waves to each other at a frequency of 1.35×10^8 Hz. (a) At what frequency does Takashi receive Yuki's radio wave? (b) At what frequency does Yuki receive Takashi's radio wave?

(a) _________ Hz

(b) _________ Hz

14.2 A spaceship is flying directly away from Earth at a speed of 0.725*c*, and the captain's family at home on Earth broadcasts their well wishes on the Family Radio System frequency of 462.5625 Mhz. To what frequency must the homesick captain tune the ship's radio to receive the broadcast? (Round your answer to three significant figures.)

_________ MHz

14.3 Starting in Toronto, Canada, a spaceship flies down Yonge Street (reportedly the longest street in the world) at a velocity of 0.330*c*. An astronaut on board measures her pulse rate at 50.0 beats per minute. (a) What is the time interval between heartbeats according to her? (b) What is the time interval between beats according to observers who are standing by the side of the road? Now, the astronaut is attached to a machine that sends a brief radio pulse for each beat of her heart back toward a monitoring device at her starting point. (c) What is the time interval measured by the monitor between consecutive radio pulses?

(a) _________ s

(b) _________ s

(c) _________ s

14.4 Albert protests the $150 ticket he received for piloting his spaceship through a red light (wavelength λ = 630 nm) on Intergalactic-5. His argument is that he was moving so quickly toward the light that it appeared to have a wavelength of λ = 546 nm, which is green. The judge agrees to dismiss the ticket, but instead fines him for speeding. The speed limit is 5.0% of *c*, with a charge of $20 for each percentage point in excess of that (for example, a speed of 0.085*c* = $70 ticket). (a) How fast was Albert traveling down the highway? Express your answer as a decimal fraction of *c*. (b) How much does he have to pay?

(a) _________ c

(b) $ _________

14.5 A police sergeant has set up a speed trap on the popular Rocketbahn route. While she is idling next to a signpost, she fires her 12.0 GHz radar gun at an oncoming sports rocket that is heading directly toward her. The radar wave hits the rocket, reflects off it, and returns directly to her detector. She detects a reflected frequency of 15.0 GHz. At what speed is the sports rocket traveling? Express your answer as a decimal fraction of *c*.

__________ c

Section 16 - Mass and energy

16.1 Find the (a) speed (as a decimal fraction of *c*) and (b) momentum of a proton that has a kinetic energy of 1000 MeV. The proton mass is 1.673×10^{-27} kg, or 938 MeV/c^2.

(a) __________ c
(b) __________ MeV/c

16.2 On the television show Star Trek®, a handheld device called a phaser vaporizes humanoid-sized bad guys, presumably turning their mass into energy, with no ill effects to the operator or to nearby objects. On the show, the energy is represented by a brief, glowing outline. In the real world, if a 50 kg alien were turned completely into energy, the energy released would be equal to how many tons of exploding TNT? The explosion of one ton of TNT releases approximately 4.0×10^{9} joules of energy.

__________ tons of TNT

16.3 In an antimatter warp drive engine, an electron traveling at a speed of 0.410*c* meets an antielectron (i.e., a positron) head on, traveling at a speed of 0.740*c*. They annihilate each other, and produce two gamma rays. What is the total energy of the resulting gamma rays? The electron and the positron have identical mass, of 9.11×10^{-31} kg, or 0.511 MeV/c^2.

__________ MeV

16.4 A neutral K meson (with mass 498 MeV/c^2), or kaon, decays at rest into two pi mesons, or pions (with mass 140 MeV/c^2 each) (a) The two pions must fly out in exactly opposite directions. Why? (b) Find the kinetic energy of each pion.

(a)
i. Conservation of momentum
ii. Conservation of energy
iii. The speed of light is always constant

(b) __________ MeV

16.5 As stated in the text, the Sun radiates 3.91×10^{26} joules per second, and thus its mass diminishes by about 4.34×10^{9} kg each second. (a) If you had a cube of water whose mass was equal to 4.34×10^{9} kg, what would be the length of one side of the cube, in km? Assume a density of 1000 kg/m^3. (b) The mass of the Sun is 1.99×10^{30} kg. Calculate how long the sun would last, in years, if it were to keep radiating at its present power level until all its mass is gone. (This is much longer than its lifetime as predicted by more sophisticated astronomical models.)

(a) __________ km
(b) __________ years

16.6 In a high energy accelerator, the Tevatron at the Fermi National Accelerator Laboratory, a proton and an antiproton moving at equal speeds make head-on collisions. Suppose that a new theory predicts that this collision can create a single particle of mass equal to 80.0 GeV/c^2. You are in charge of running the initial experiment. The proton and antiproton each have equal rest masses of 938.38 MeV/c^2. What minimum proton/antiproton speed is necessary for the new particle to be created? Express your answer as a decimal fraction of *c* to five significant figures (or else you will just find that $v = c$).

__________ c

16.7 The Stanford Linear Accelerator is capable of producing beams containing 30 GeV electrons. (That is, the total energy of each electron is 30 billion electron volts.) What is the Lorentz factor for these electrons?

____

16.8 In a high energy accelerator, protons and antiprotons (with equal rest masses of 938 MeV/c^2), each with kinetic energy equal to 2.00 GeV (1 GeV = 1000 MeV), make head-on collisions. The two particles annihilate each other to form a K− and a K+ meson, each having rest mass equal to 494 MeV/c^2. (a) What is the speed of a proton before the collision, as measured in the lab? Express your answer as a decimal fraction of *c*. (b) What is the kinetic energy of each K meson?

(a) __________ c
(b) __________ MeV

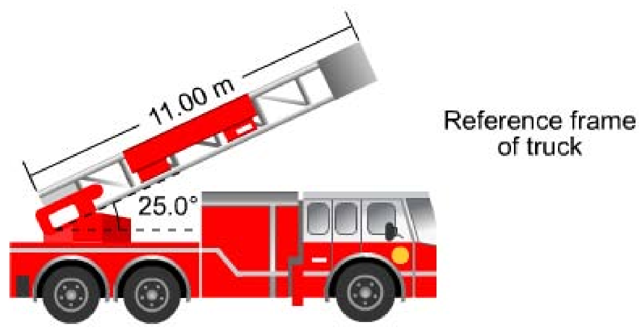
Additional Problems

A.1 A runner carrying a horizontal pole of proper length 10 m runs at 0.866*c* through a barn whose proper length is 5.0 m. The farmer measures the length of the pole as only 5.0 m, due to length contraction. He is then determined to trap the runner in the barn by slamming the front door shut behind her when the pole is completely inside. The pole will smash through the back wall, but for an instant, the entire pole will be inside the barn. On the other hand, the runner observes that the barn's contracted length is only 2.5 m and declares that there is no way her 10 m pole is going to fit inside the barn. (a) Who is right? (b) Explain your answer.

(a) i. The farmer is right
ii. The runner is right
iii. They are both right
iv. They are both wrong

(b)

A.2 An 11.0 m long ladder atop a fire truck is pivoted until it is angled at 25° above the horizontal, as measured by a firefighter riding in the truck. An observer in the street sees the fire truck race by at a speed of 0.775*c*. (a) How long is the ladder, according to her? (b) At what angle to the horizontal is it oriented, according to her?



(a) __________ m
(b) __________ °

A.3 This question explores the relativistic basis for the existence of magnetic fields. In a current-carrying conductor such as a copper wire, only the electrons actually move, but here, for the sake of clarity, we will consider a symmetrical model where the current is carried by both positive and negative charges. Imagine a long stream of positive charges, separated by a distance *d*, moving to the right along the *x* axis with speed v_0. Superimposed on this is a stream of negative charges of the same magnitude, also separated by *d*, but moving to the *left* along the *x* axis at the same speed. (a) In the laboratory frame of reference, is there an electric field at the location (0.0, −1.0) cm? (b) In the laboratory frame, there is a magnetic field at the location (0.0, −1.0) cm. What is its direction? (c) Let a positive test charge move toward the right with speed *v*, one centimeter below the *x* axis. At the point (0.0, −1.0) cm, what is the direction of the force it experiences? (d) Now consider the reference frame *S ′* of the moving particle. There is a magnetic field at the location of the charge, but does the test particle feel any force due to the magnetic field? Hint: in *S ′*, the particle is stationary. (e) Explain the apparent contradictory difference in the force acting on the test charge in parts c and d.

(a) ○ Yes ○ No

(b) i. Positive x
ii. Negative x
iii. Positive y
iv. Negative y
v. Toward you
vi. Away from you

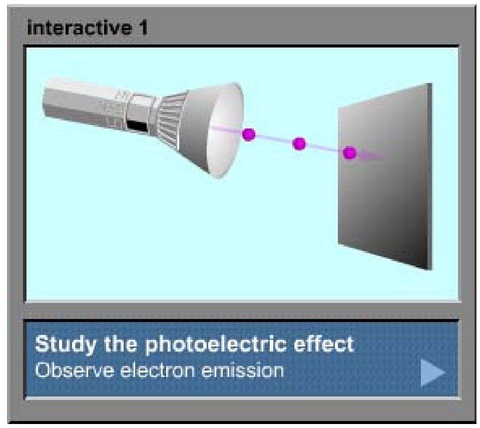
(c) i. Positive x
ii. Negative x
iii. Positive y
iv. Negative y
v. Toward you
vi. Away from you

(d) ○ Yes ○ No

(e)

Chapter 41 Quantum Physics Part One

41.0 - Introduction



Quantum physics is the branch of science required to fully explain the behavior of light, its interaction with matter, and the behavior of exceedingly small particles such as atoms and electrons. As scientists began to discover in the late 19th and early 20th centuries, certain principles and techniques of classical physics fail utterly when applied to light and to atomic-scale systems.

Although “quantum physics” often connotes mystery and difficulty, its applications are very real. You may be pleasantly surprised at how much you can understand when you are equipped with just a few fundamental concepts from this science.

In particular, you can learn the principles governing the functioning of two of the most pivotal technologies of the last half-century: semiconductors and lasers.

Why are these two technologies so important? Without semiconductors, there would be neither transistors, nor the microprocessors built from them. Semiconductor-based microprocessors serve as the “brains” of computers and are found in digital cameras, cell phones, and automobiles: wherever engineers want “smart” behavior. Semiconductors are also used in various types of computer memory, such as random access memory (RAM). Semiconductor chips not only “think,” they also “remember.”

In recent years, **connecting** all these semiconductor devices has become the central thrust of the information processing industry. The Internet, cell-phone “fixed-rate calling plans”, video on demand, downloadable music, and even the Web-based version of the textbook you are now reading all rely on the cheap and rapid transmission of information over wired or wireless networks. The two technologies most responsible for creating this networking revolution of rapidly decreasing costs and dramatically increasing bandwidth have been the microprocessor and the laser.

How do these two technologies enable networking? Communication networks use devices like *routers* and *high-speed switches* to transmit data. These devices rely on microprocessors to determine where to send their information and they form parts of extended physical systems that use lasers to move the information at light speed over fiber-optic cables.

Lasers give you access to data from sources both distant and nearby. In addition to sending data around the world, they are used on your desktop or in your home to read the data stored on CDs and DVDs (not to mention their use in stores to read data codes on your purchases). Without lasers, vinyl records and “floppy disks” might still be the primary means of storing audio and digital data. Believe us: If you have never used a floppy disk, you haven’t missed much.

How does quantum physics relate to the working of these devices?

Explaining what is meant by “quantum” is the place to start. A key tenet of quantum physics is that particles in some systems, like the electrons in hydrogen atoms, exist only at certain energy levels. Physicists say the energy levels of the electrons in an atom are *quantized*.

It may be easiest to explain a quantum property by first considering its opposite, a property that is *continuous*. Consider the potential energy of a bucket that is raised or lowered by a rope. You can raise it 1.000 meters off the ground, or 1.001 m, or 1.002 m, or however much you like. By controlling its height, you can make its potential energy whatever you like. The range of possible energy values is continuous: say 10.00 joules, 10.000017 J, 10.027 J and so forth.

Electrons prove not to be as flexible. The electrons around a hydrogen (or other) atom exist only in states with certain discrete energy values; for instance, two possible energy levels for the hydrogen electron are −1.51 eV (electron volts) or −3.40 eV (-2.42×10^{-19} J or -5.44×10^{-19} J respectively). Between these two values lies a forbidden gap, and a hydrogen atom’s electron is never observed with energies in that range. Physicists say an electron’s energy is quantized, that it only exists at certain levels. In the example of the hydrogen electron mentioned above, you will **never** observe an energy of −1.6 eV or −2.9 eV, since those are forbidden.

You may ask: How does an electron change between energy levels? How can it “move” across a “forbidden gap” to a higher or lower energy state, if intermediate energy values are forbidden? You may not be satisfied with the answer, but it is most straightforward to say: Those are simply the only values that have ever been measured. Any time a scientist measures a property of an atomic electron (such as its energy, or its angular momentum), she only observes results from a particular set of values that can be predicted with extreme accuracy by quantum physics. It is impossible to “catch” an electron in any in-between state.

The simulation at the right reproduces one of the key experiments that led to the widespread acceptance of the ideas of quantum physics. Einstein explained data from a more sophisticated version of the same experiment, known as the photoelectric effect, to earn his Nobel Prize.

Describing the experiment is simple. Scientists had noticed that when they shined a beam of light on a metal, electrons were released from the illuminated surface. From their perspective, this was not particularly surprising: The energy of the light was transferred to the electrons, allowing some of them to escape their bonds to the atoms of the metal.

The emission of electrons could be explained by classical physics. Light was a wave with energy, and that energy could provide a “kick” to electrons as the metal absorbed the light.

Although the experiment and the expected outcome are simple in concept, the detailed results were quite surprising and could **not** be explained by classical physics. Some colors of light, such as red, could not eject electrons from the surface, no matter how bright the beam was. On the other hand, other colors, such as violet, were effective at ejecting electrons from the metal surface, even when the beam intensity was very low. It was the frequency of the light, and not its intensity, that determined whether or not electrons were ejected.

What the scientists observed does not make sense if light is conceived of solely as a wave. Let's compare their observations in terms of water waves crashing against a wooden dock. It is as if low-frequency waves (with their crests arriving, say, every five seconds) could **never** rattle the dock enough to knock free the timbers that make it up, even if they were giant waves 50 meters tall.

Now imagine centimeter-high waves arriving more frequently, say every second. Imagine that these small but frequent waves could knock loose pieces of the wood from the dock. Water waves with these effects would be confusing to observe, and you might be as confused as the scientists who observed dim but high frequency light freeing electrons from samples of metal.

Einstein successfully explained the photoelectric effect. He argued that light has both a wave nature and a particle nature. Electromagnetic radiation, he said, consists of small packets called photons. Photons are small "chunks" of light energy. More energetic light consists of more photons, **not** larger, more energetic waves. In addition, he stated that the higher the frequency of light, the more energetic the photons that make up the light. Dim but high frequency light can eject electrons because of the interaction between the energetic photons that make it up and the atoms of the metal. It is the energy of the individual photons that matters, not the overall energy of the light.

Conceiving of light as consisting of photons could explain another mysterious result: More intense light of a certain color caused more electrons to be emitted, but their maximum kinetic energy was exactly the same. The classical physicist would expect the "larger wave" of the more intense light to cause higher-energy electrons to be released, but this did not happen. Einstein's theory explains why: more intense light consists of more photons, each with the same energy as before. Again, it is the interaction between an individual photon and an individual atom that matters.

You just read a brief summary of some crucial points in quantum physics. You will become familiar with the photoelectric effect by using the simulation on the right. In your experiment with this effect, the flashlight can shine red or violet light. It can be set to emit either low or high intensity light.

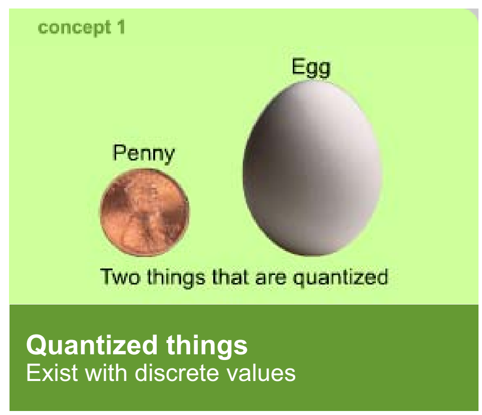
When you press GO, you will see photons moving in slow motion from the flashlight toward the metal. When appropriate, we show electrons escaping from the metal.

Start the simulation with the light set to LOW. One color of light will cause electrons to escape the metal being used in our simulation; another will not. Red light has a longer wavelength but a lower frequency than violet light. Which of the two colors of light do you think will cause electrons to be emitted?

Now set the intensity of the light to HIGH, and try both colors again. What do you expect will change when you make this change? What do you think will stay the same?

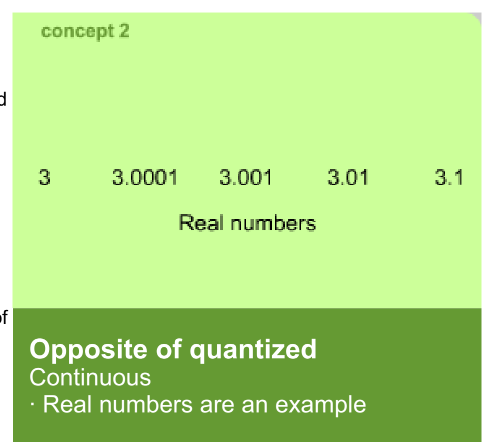
41.1 - Quantum

Quantum: The smallest amount of something that can exist independently.



Quantum refers to the fundamental or least amount of something. For instance, the quantum of money in the U.S. is the penny. Your net worth will be a multiple of that quantum. You can be a pauper worth one penny, a millionaire with a worth of 100,000,000 pennies, or a starving college student with a net worth of −5,012 pennies. However, you cannot legally use three-quarters of a penny, or 1.45 pennies, or $3\pi/4$ pennies. There are many similar examples of things that come in discrete amounts: the number of siblings you have, the number of eggs you can purchase at a store, and so on.

A physicist would say that things like money or siblings or eggs are *quantized*. Calling something quantized means that it is grainy; it is the opposite of continuous. Using the example mentioned in this chapter's introduction, one would say that the height and energy of a bucket being raised by a rope are continuous quantities, as are the height and energy of an elevator car. In contrast, the height and energy of elevator **stops** are quantized; they occur solely at discrete points. You only see buttons for the first, second and third floors, not for the 1.75th floor.



A mathematical example of something that is continuous is shown on the right: real numbers. Examples of real numbers are 3, or 3.1, 3.01, 3.001, 3.002 and so forth. The set of real numbers is continuous, **not** quantized.

Although the idea of quantization may seem intuitive for money, it is much less obvious in some areas of physics: Scientists now know that many things once thought to be continuous are in fact quantized. Albert Einstein, for instance, showed that the energy of any precise color of light is quantized.

Prior to Einstein scientists expected properties of light, such as its energy, to be continuous. Why? In the 18th and 19th centuries, a series of discoveries had led most scientists to conclude that light was a wave. They knew very well that the energy of a mechanical wave is continuous, not quantized. An ocean wave, for instance, can have a height (amplitude) of 1.01 meters, or 1.04 meters, or any value in between, and its energy will depend on that amplitude. Since light was believed to be a wave, scientists concluded that its energy would be continuous as well, and that for example, they could create a beam of a certain frequency of blue light with any desired energy simply by making the light brighter or darker.

However, as the next sections discuss, in the early 20th century it became increasingly clear that light is quantized: It consists of small chunks or packets of energy. The energy of a beam of a particular color of light must be a multiple of the energy of the packets that make it up. This realization had profound implications for the understanding of both light itself and the atoms that emit the light.

41.2 - Balmer series

The electrically excited neon gas in this sign emits light at several sharply defined red-orange wavelengths.

“The most important result of the application of quantum mechanics to the description of electrons in a solid is that the allowed energy levels of electrons will be grouped into *bands*.” So wrote Andy Grove, then an employee of the Intel Corporation and a faculty member of the University of California at Berkeley, in his text on the physics and technology of semiconductor devices. Grove became the chairman of Intel during its rise to power, prestige and profitability.

Grove cites two quantum principles. First, the concept that there are “energy levels” for electrons, and second, that these are grouped into bands (the emphasis in the quote is his).

For now, we will simplify our discussion by focusing on energy levels. Bands refer to the fact that certain electrons exist at energy levels that are close to one another. In practical applications, the distinction between a band and an energy level is often dropped.

Grove’s words convey how important the ideas of quantum theory are for semiconductor science and technology. We use these words to motivate the next few sections of this book. One might wonder: How did physicists discover that atomic electrons had discrete energy levels? In other words, how did they first learn that the energy levels of electrons are quantized, not continuous?

Physicists advanced the theory as their observations forced them to. The story begins in 1666 when Newton showed that a prism could disperse sunlight into a spectrum of colors. Today, one would say Newton showed that sunlight comprises light of many wavelengths: Light perceived as white is in fact made up of a rainbow of components of various hues. When it was first discovered, the spectrum of sunlight seemed to be a continuous gamut of colors.

A series of later experiments convinced scientists that light had a wavelike nature. They could create interference patterns with light that were conceptually identical to patterns created by water waves. Physicists even found that they could measure the wavelengths of various colors of light. One mystery of science seemed to be solved: Light was a wave.

However, in 1814, the German physicist Joseph von Fraunhofer made careful observations using a thin slit, and discovered that the spectrum of sunlight contained many narrow dark lines, or gaps. In other words, certain wavelengths of light were not present in the spectrum he was observing. He discovered that the spectrum of sunlight was not continuous.

Throughout the 1800s, scientists studied the light emitted and absorbed by various gases. They discovered that a gas like hydrogen only emits or absorbs light of specific wavelengths. By 1880, the wavelengths of the *spectral lines* of various elements, including most famously hydrogen, were well known. In Concept 1 you see an illustration of the spectral lines in the *emission spectrum* of excited hydrogen gas.

The distinct colors and wavelengths of light you see are characteristic of light emitted by this element. Similar lines, but with different colors − wavelengths − can be found when the light of a neon sign, or the glow of a fluorescing ruby, is analyzed. (Each element has a corresponding *absorbtion spectrum*, consisting of dark lines at exactly the same wavelengths, against a rainbow background.)

As you can see, the spectral lines of hydrogen are sharp and distinct, not blurred. For instance, the red light you see has a wavelength of 656.3 nanometers, the blue-green light has a wavelength of 486.1 nm, and the violet light has a wavelength of 434.1 nm. Hydrogen atoms emit this light after being “excited” by an electrical discharge through the gas, caused by a potential difference of 5000 volts applied between two electrodes.

The discrete nature of the hydrogen spectrum puzzled and intrigued physicists. Why did the light emitted by hydrogen only exist at certain wavelengths, rather than being continuous like a rainbow? And why at these particular wavelengths? Was there any way to predict the wavelengths?

A Swiss high school teacher, J. J. Balmer, analyzed the pattern. He determined that the wavelengths were not random, but could be determined using the formula in Equation 1. The constant R_H that appears in Balmer’s formula is called the *Rydberg constant*.

concept 1

Hydrogen emission spectrum
Is not continuous

equation 1

J. J. Balmer (1825 - 1898)

Balmer series

$$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \text{ for } n = 3, 4, 5, \ldots$$

λ = wavelength
R_H = Rydberg constant
n = integer
Constant $R_H = 1.097\ 37\times10^7\ \text{m}^{-1}$

example 1

What is the lowest visible frequency of light emitted by

Intriguing mathematical patterns did not stop with the set of spectral lines known as the *Balmer series*. There are also wavelengths emitted by hydrogen that lie outside of the visible spectrum, which are predicted by formulas very similar to the one in Equation 1.

Later scientists determined that similar relationships existed for the spectral lines of other elements, as stated by the *Rydberg-Ritz combination principle*. Although the hydrogen atom is often used to discuss this principle in order to keep things simple, it applies to all atoms. For instance, the neon sign you see in the photograph above is displaying red light at the wavelengths of several of its spectral lines.

The physicists who determined these mathematical relationships did not know **why** the wavelengths of the spectral lines followed the patterns they did. These data were just too far ahead of the theory of atomic structure. Scientists could observe the discrete spectral lines of the emitted and absorbed light, and note the mathematical relations that predicted their wavelengths, but could only speculate as to the cause.

However, the work was underway. The light emitted by hydrogen was found to have a discrete spectrum, one that could be predicted by a formula. It was a tantalizing clue about the quantized nature of atoms.

hydrogen?

$$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$$

$$c = f\lambda \text{ so } \frac{1}{\lambda} = \frac{f}{c}$$

$$\frac{f}{c} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$$

$$f = cR_H\left(\frac{1}{2^2} - \frac{1}{3^2}\right) = cR_H\left(\frac{5}{36}\right)$$

$$f = \tfrac{5}{36}(3.00\times10^8\,\tfrac{\text{m}}{\text{s}})(1.097\times10^7\,\tfrac{1}{\text{m}})$$

$$f = 4.57\times10^{14}\text{ Hz}$$

41.3 - Planck and blackbody radiation

Molten metal. The hottest parts (white) emit a full spectrum of light. Cooler parts emit narrower spectra centered on yellow, orange, or red.

While some scientists were studying the spectral lines of gases, others were plunging into other puzzles in the field of electromagnetic radiation. For instance, Max Planck was studying a form of radiation called *blackbody radiation*.

Although a blackbody is a theoretical construct, the basic phenomenon will be familiar to any clay sculptor. When he places a piece in a kiln to fire it, the clay will be gray. As the clay gets hotter and hotter, it begins to glow. The color of the glow will change from orange to yellow to white as the clay increases in temperature. Planck was trying to answer a simple question: Why does the color of the light change with temperature?

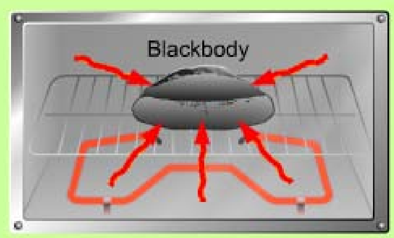
He used the concept of a blackbody. A *blackbody* absorbs all the electromagnetic radiation (such as light) that strikes it, not reflecting any. It reaches thermal equilibrium by re-emitting the energy as a spectrum of radiation called blackbody radiation. The amount of energy it radiates at each frequency in this spectrum depends only on its absolute (Kelvin) temperature. The physicists of Planck's era believed that they should be able to formulate a theory to predict the overall perceived color, or radiation spectrum, of a blackbody at any temperature by using existing theories.

Physicists recognized however, that the theory that Planck inherited had a major flaw: It predicted that for any temperature, the amount of energy radiated would be unlimited at higher frequencies. The model that Planck would have studied in school predicted that if a blackbody were heated in an oven, it would radiate an infinite amount of energy in the high-frequency end of the electromagnetic spectrum. This (obviously incorrect) prediction had a lurid name, the *ultraviolet catastrophe*. Since scientists knew that the amount of energy radiated was finite and not unlimited, they knew that it was their theory that was the catastrophe, so to speak.

Planck found that he could avoid the catastrophic disagreement between observation and prediction by assuming that the radiation absorbed or emitted by a blackbody is quantized. He stated that the energy absorbed or emitted must obey the relationship shown in Equation 1. The energy of the radiation at a given frequency is an integer multiple of a constant h (*Planck's constant*) times the frequency of the radiation. Planck's constant is $6.626\,08\times10^{-34}$ J·s.

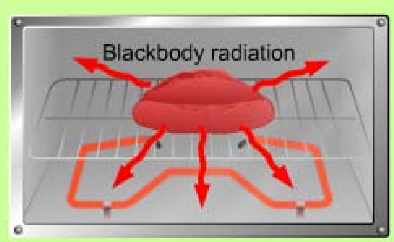
Planck's work was a crucial step in the development of quantum theory: He showed how a mystery of physics could be resolved by applying the idea of quanta to the emission and absorption of electromagnetic radiation. As you will see shortly, Einstein understood this development, realized its importance, and took the next crucial step. In essence, Einstein showed how one should think about that "n" in front of the "hf."

concept 1



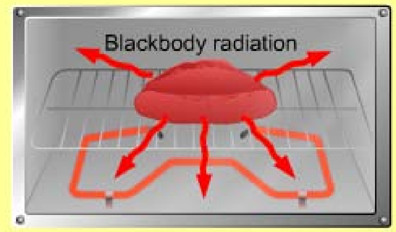
Blackbody absorption
Blackbody absorbs all electromagnetic radiation

concept 2



Blackbody radiation
Only quantized emission correctly predicts radiation spectrum

equation 1



Energy of blackbody radiation

$E = nhf$

E = energy, n = a positive integer
h = Planck's constant
f = frequency
Constant $h = 6.63\times10^{-34}$ J·s

41.4 - Photons

Photon: A packet of light. The fundamental unit or quantum of light.

Max Planck showed how the radiation emitted by an ideal object in thermal equilibrium with its surroundings could be explained if the radiation emitted or absorbed by the body was quantized. His theory had a somewhat marginal existence for four years until Albert Einstein, then employed as a Swiss patent clerk, began to consider it in depth.

Planck's theory stated that the energy (radiation) **absorbed or emitted by an object** had to be taken up or released in discrete chunks, as quanta. However, he still conceived of the radiation itself as a wave. His theory simply stated that matter absorbed or emitted the radiation in discrete amounts.

Einstein took the next bold step: He stated that the radiation itself was quantized. He saw that what Planck had been studying was not just how matter absorbed and emitted radiation, but the basic nature of the radiation itself.

Einstein stated that light of any frequency is quantized in units now called *photons*. The energy of a photon is proportional to its frequency f.

He wrote, "The energy in a beam of light is not distributed continuously through space, but consists of a finite number of energy quanta, which are localized at points, which cannot be subdivided, and which are absorbed or emitted only as whole units." (This was in the same year that he published his special theory of relativity; not a bad year.)

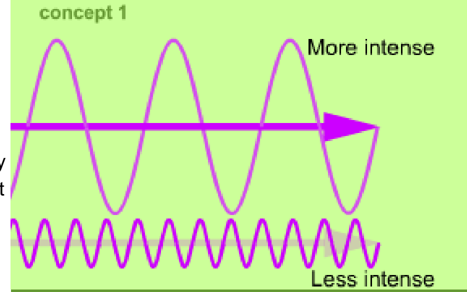
This new model challenged the previous concept that light behaved solely as a wave. Instead, it stated that light could also be conceived of as a stream of packets of energy, almost as particles.

The energy of each photon could be calculated with the equation $E = hf$. In other words, red light of frequency 4.60×10^{14} Hz cannot have just any energy level; its energy is always an integer multiple of hf. One photon of red light of frequency 4.60×10^{14} cycles per second has 3.05×10^{-19} J of energy, two photons of this frequency red light have 6.10×10^{-19} J of energy, and there is no such thing as 1.5 photons of red light, any more than there can be 1.5 electrons.

It was not easy for scientists to accept Einstein's new theory. When four elite scientists − including Planck − nominated Einstein to the Prussian Academy of Science they wrote, "That he may have missed the target in his speculations, as, for example, in his hypothesis of light quanta, cannot really be held too much against him, for it is not possible to introduce fundamentally new ideas, even in the most exact sciences, without occasionally taking a risk."

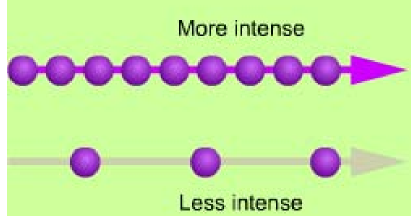
One year later, the American physicist Robert A. Millikan reported a precise confirmation of Einstein's equation for the energy of the photon, $E = hf$.

concept 1



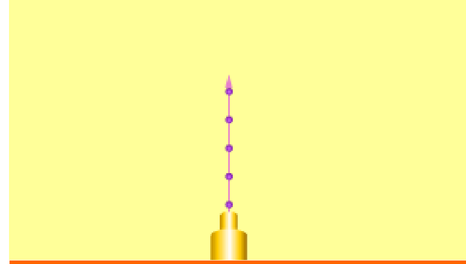
Pre-Einstein conception of light
Light is a wave
· Energy can vary continuously
· Einstein demonstrated flaws in the wave model of light

concept 2



Photon
Packet of energy
· Light is quantized
· Brighter light = more photons

equation 1

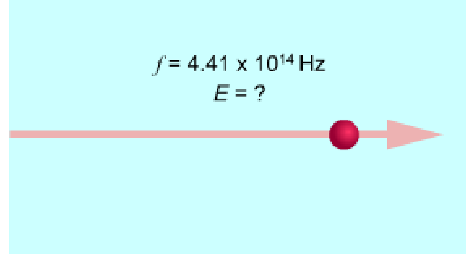


Energy of a photon

$E = hf$

E = energy
h = Planck's constant
f = frequency
Constant $h = 6.63\times10^{-34}$ J·s

example 1



What is the energy of this photon of red light?

$E = hf$

$E = (6.63\times10^{-34}\text{ J·s})(4.41\times10^{14}\text{ s}^{-1})$

$E = 2.92\times10^{-19}$ J

41.5 - Sample problem: solar radiance

Take the "typical" wavelength of light to be 550 nm. What is the energy of a photon of that light? If you were to stare into the Sun (a very bad idea!), how many photons per second would enter one of your eyes? Use 7.85×10^{-7} m^2 for the surface area of the pupil and assume that the intensity of sunlight on the Earth's surface at your location is 1000 W/m^2.

Variables

total energy	E
energy of a single photon	E_p
frequency of light	f
wavelength of light	$\lambda = 550\times10^{-9}$ m
Planck's constant	$h = 6.63\times10^{-34}$ J·s
intensity of light	$I = 1000$ W/m^2
surface area of one pupil	$A = 7.85\times10^{-7}$ m^2
speed of light	$c = 3.00\times10^{8}$ m/s
number of photons per second	N

What is the strategy?

1. Determine the energy of each photon by using Einstein's equation for the energy of a photon. To determine the photon's frequency, use the relationship between the speed of light, its wavelength and its frequency.
2. Calculate the power of the light entering your eye by using the relationship of intensity, power and surface area.
3. Compare the power of sunlight hitting your eye to the energy of a single photon and calculate how many photons are hitting your eye each second.

Physics principles and equations

The energy of a photon is

$$E = hf$$

The relationship between wave speed (in this case, c), wavelength and frequency is

$$c = \lambda f$$

Intensity can be calculated as

$$I = \frac{P}{A}$$

Power equals

$$P = \frac{\Delta E}{\Delta t}$$

Step-by-step solution

We first compute the energy of a single incident photon, answering the first part of the question above.

Step	Reason
1. $E_p = hf$	energy of a photon
2. $c = \lambda f$	wave speed, wavelength and frequency
3. $f = \frac{c}{\lambda}$	solve for frequency
4. $E_p = \frac{hc}{\lambda}$	substitute step 3 into step 1
5. $E_p = \frac{(6.63\times10^{-34}\ \mathrm{J\cdot s})(3.00\times10^{8}\ \frac{\mathrm{m}}{\mathrm{s}})}{(550\times10^{-9}\ \mathrm{m})}$ $E_p = 3.62\times10^{-19}$ J	evaluate

Next we calculate the power of the sunlight entering your pupil. To do so, we use the intensity equation above, the given intensity of the light and the area of the pupil.

Step	Reason
6. $I = \frac{P}{A}$	intensity
7. $P = IA$	solve for power
8. $P = (1000\ \text{W/m}^2)(7.85 \times 10^{-7}\ \text{m}^2)$ $P = 7.85 \times 10^{-4}\ \text{W}$	evaluate

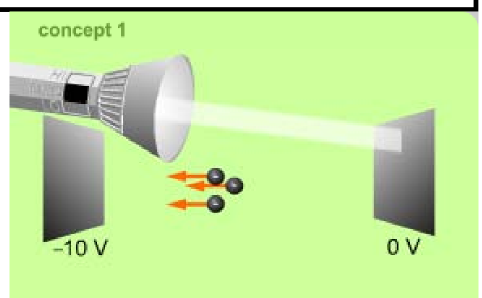
We have determined the power of the light entering your pupil and the energy of a single photon. Now we will relate these two values to find the number of photons per second that enter your eye.

Step	Reason
9. $P = \frac{\Delta E}{\Delta t}$	definition of power
10. $P = NE_p$	power and photon rate
11. $N = \frac{P}{E_p}$	solve equation 10 for N
12. $N = \frac{(7.85 \times 10^{-4}\ \text{W})}{(3.62 \times 10^{-19}\ \text{J})}$ $N = 2.17 \times 10^{15}$ photons per second	evaluate

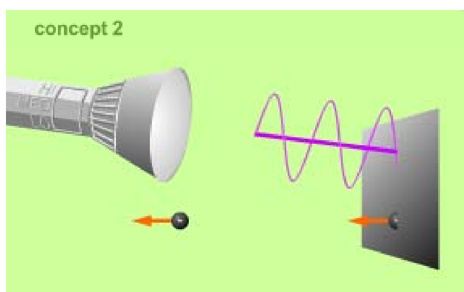
Note that N is a very large number of photons per second. Sensors in the eye's retina can actually respond to just a single photon. A single molecule in one "rod" cell in your eye can absorb one photon, triggering a chemical reaction that sends a signal to the optic nerve. However, neural filters only let a signal go to the brain if approximately five to nine photons arrive every 100 ms. This limit prevents the visual fuzziness or "noise" that would exist under conditions of low intensity light if the eyes were too sensitive.

41.6 - Photoelectric effect

Photoelectric effect: The ejection of electrons from a material due to light striking it. Aspects of this effect were used by Einstein to demonstrate the quantization of light.



Light and electrons
Shining light on metal causes it to emit electrons



Incorrect expectation

What caused Einstein to believe that light was quantized? In the year 1905 he used a quantum model of light to explain the results of an experiment that could not be explained using classical electromagnetic theory. In fact, Einstein won the 1921 Nobel Prize in Physics "for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect."

In 1887, Heinrich Hertz had shown that shining light on metal could cause electrons to be ejected from the metal. You can think of this process as analogous to evaporation. When light shines on water, it can cause some of the water molecules to escape as a gas. When light shines on metal, some of the electrons can escape the metal.

We illustrate this phenomenon in Concept 1, with electrons escaping from the metal electrode on the right. The "ejected" electrons could be readily explained by the classical model of light as a wave composed of electric and magnetic fields. These waves transported energy, and it made sense that some of the electrons in the metal could absorb enough of this energy to escape the attractive force binding them to atoms in the metal.

In Concept 1, you also see parts of an apparatus used to conduct experiments whose results were not so readily explained using classical theory. It consists of two electrodes enclosed in a vacuum. The left electrode has a lower electric potential than the right electrode, which means it will repel electrons. You may also think of the apparatus in this way: an electric field is established between the electrodes that points from the right to the left. This field "pushes" the electrons to the right, back toward the electrode they have escaped from.

Shining light on the right-hand electrode causes electrons to be ejected from it and to move to the left. The faster they are moving, the more negative the electric potential of the left electrode has to be to keep them from reaching it. By adjusting the electric potential on the left electrode so that the electrons "just fail" to reach it, an experimenter can determine the maximum kinetic energy of the electrons.

An expected result of this experiment would be that, the more intense the light shining on the right electrode, the more energy its electrons would absorb, and the faster they would move when ejected from the metal. It would be like chopping wood with an ax. As you chop the wood, the harder you strike, the faster you expect some of the chips to fly off.

However, to the surprise of the experimenters, this proved **not** to be the case. There is **no** correlation between the intensity of the light and the maximum kinetic energy of the electrons. Rather, when the intensity of the light is increased, more electrons are emitted. Instead of the kinetic energy of the escaping electrons increasing, only their number increases. The incorrect "classical" expectation is shown in Concept 2, and the actual observed behavior in the experiment is shown in Concept 3 (refresh your browser to restart these animations). In terms of the ax analogy: It is as if hitting the log harder results not in more energetic chips, but in more chips of the same energy flying off.

Another surprising result was that when the frequency of the light was below a certain value, known as the *cutoff frequency*, the light could be of great intensity, but no electrons at all would be ejected. Classical physics cannot explain these phenomena. To use the ax example one more time: If you strike a log very infrequently, but with great force (energy), you would expect chips to fly off. However, if the photoelectric effect applied to wood chopping, then when you chop at a slow rate, no chips will fly off, no matter how hard you strike the log. Very odd!

Quantum theory, however, explains both these results quite neatly. Considering light as packets of energy means that with more intense light more packets are striking the metal each second. The energy of each packet does not change with intensity; the number of packets does. More photons (each having the same energy) striking the right electrode each second increases the rate at which electrons are ejected from the metal, but not their maximum kinetic energy.

This can also be likened to knocking over bottles with bullets from a rifle. In classical theory, more "intensity" means switching to a bigger, more powerful rifle, so that each bullet will hit a bottle with more energy. The same number of bottles will still be hit; they just go flying off the target faster.

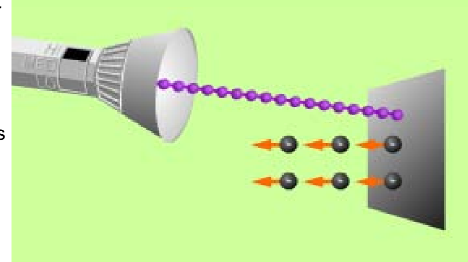
In quantum theory, more "intensity" does not mean switching to a larger caliber, more powerful rifle. More intensity means firing more bullets per second from the same rifle, which means more bottles get knocked over. The increased intensity does not involve changing the energy of the bullets, so the bottles always get struck with the same amount of energy.

Quantum theory also explains why low-frequency light cannot cause any electrons at all to be emitted. The energy of each photon equals hf. Using intense light of low frequency means that many low-energy packets of light are striking the metal, but no individual packet has enough energy to raise an electron to an energy level high enough for it to escape.

The minimum energy an electron needs to escape the metal is called the *work function*, and low-frequency photons have less energy than this. To use the rifle analogy: Employing low-frequency light is like changing from bullets to spitballs. No spitball alone has enough energy to knock over a target. In quantum theory, increasing the intensity of light means the number of spitballs increases, **not** the energy of a spitball. Just as no spitball is energetic enough to knock over a bottle, no low-frequency photon in the experimental apparatus is energetic enough to free an electron.

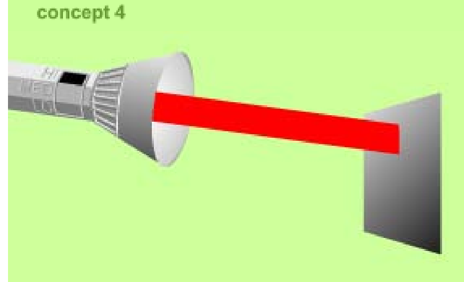
Speed/energy of an electron corresponds to light's intensity??

concept 3



What was actually observed
Number of electrons increases with light's intensity

concept 4



Relationship of frequency and electrons
No electrons emitted
· For colors below a certain frequency
· No matter how intense the light

concept 5

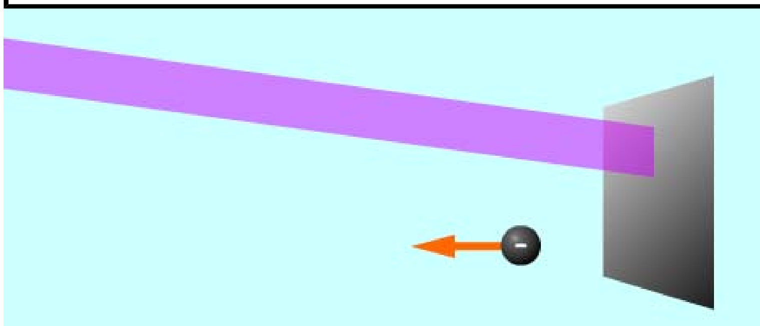
Quantum theory (and Einstein) to the rescue
More intense light → more photons
More photons → more ejected electrons

Frequency/energy of light and electrons

Lower frequency light → photons less energetic

Less energetic photons cannot free electrons

41.7 - Sample problem: photoelectric effect



Ultraviolet light with a frequency of 1.03×10^{15} Hz is incident on a metallic sodium surface. The work function, the minimum amount of energy required for an electron to be ejected from this material, is φ = 2.36 eV. What is the maximum kinetic energy an escaping electron can have?

The electrons in a particular metal have a range of different energies and require various amounts of additional energy in order to be "freed" when the metal is struck by a photon.

The *work function* of a material is the least amount of energy required to release any electron from it. In other words, it is the amount of energy required to free the "least attached" electron.

Variables

energy of photon	E
maximum kinetic energy of electron	KE_{max}
work function for sodium	$\varphi = 2.36$ eV
Planck's constant	$h = 6.63\times10^{-34}$ J·s
frequency of ultraviolet light	$f = 1.03\times10^{15}$ Hz

What is the strategy?

1. Use the conservation of energy. This means that the energy of the incoming photon must equal the energy required to free the electron plus the kinetic energy of the electron as it flies away from the surface. (Ultraviolet light does have enough energy to free an electron in this case.)
2. Observe that the electrons that are least bound to the metal will be the electrons that leave it with the maximum kinetic energy. The less energy that is "spent" by a photon to free an electron, the more energy there is "left" that can go to increasing the electron's kinetic energy.

Physics principles and equations

Energy is conserved.

The energy of a photon is

$$E = hf$$

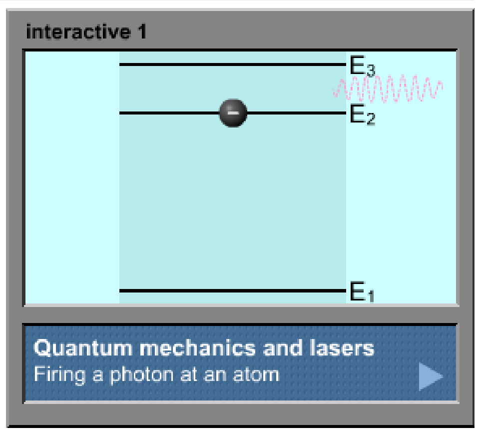
Step-by-step solution

Step	Reason
1. $E = KE_{max} + \varphi$	conservation of energy
2. $KE_{max} = E - \varphi$	solve for KE_{max}
3. $E = hf$	energy of a photon
4. $KE_{max} = hf - \varphi$	substitute equation 3 into equation 2
5. $hf = \left(6.63 \times 10^{-34}\ \frac{\text{J}}{\text{s}}\right)\left(1.03 \times 10^{15}\ \text{s}^{-1}\right)$ $hf = 6.83 \times 10^{-19}$ J	photon energy
6. $hf = 6.83 \times 10^{-19}\ \text{J}\left(\frac{1\ \text{eV}}{1.6 \times 10^{-19}\ \text{J}}\right)$ $hf = 4.27$ eV	convert units
7. $KE_{max} = 4.27\ \text{eV} - 2.36\ \text{eV}$ $KE_{max} = 1.89$ eV	evaluate

We converted to electron volts in step six because in quantum physics, energies are commonly stated in electron volts rather than joules. An electron with kinetic energy $KE_{max} = 1.89$ eV would be moving at a speed of 8.19×10^5 m/s, which is approximately 0.27% of the speed of light. At this speed, relativistic effects do not play a significant role and are ignored in the solution above.

The equation derived above, $KE_{max} = hf - \varphi$, was an important result that was used to confirm the particle-like nature of light. Einstein first stated it in 1905, and Robert A. Millikan experimentally verified it shortly thereafter.

41.8 - Interactive problem: photoelectric effect



Laser engineers use the principles of quantum physics. In the simulation on the right, you will experiment with one of the fundamental phenomena that underlie the working of a ruby laser. Rubies are crystals of aluminum oxide with chromium impurities that give them their distinctive red color. More importantly for engineers, chromium is primarily responsible for ruby's lasing properties.

The simulation focuses on one of the outermost electrons of a chromium atom. The initial step to make a laser work is to excite the electron so that it is in a high energy state. The electron starts at a low energy state called E_1. Your first goal is to increase this electron's energy, and cause it to jump to a higher energy state called E_3.

How can you boost the energy of the electron from E_1 to E_3? Here, you may fire at the atom either photons of red light with a frequency of 4.32×10^{14} Hz, or photons of green light with a frequency of 5.45×10^{14} Hz. One photon of the red light used in the simulation has an energy of 1.79 eV, and a photon of the green light has an energy of 2.26 eV. These photon energies are calculated using the formula $E = hf$, where h is Planck's constant and f is the frequency of the light.

Conduct a few experiments by selecting photons of different colors (energies) and then pressing GO. Do photons of both frequencies have an effect on the outermost electron of the chromium atom, or just one frequency? After testing the photons, can you deduce what the energy difference is between E_1 and E_3?

After you fire the appropriate photon, the electron will have energy E_3, but this energy level is highly unstable. The electron will rapidly and spontaneously fall to an intermediate energy level E_2. This new state is relatively stable compared to the higher state, and we will pause time in the simulation when it reaches this "metastable" state. (A photon is emitted when the electron drops from E_3 to E_2 but that photon is irrelevant to the operation of the laser.)

Your final goal is to stimulate the chromium atom to emit another photon, which it will do when the electron drops from energy E_2 to E_1. You do this by firing a new photon at the chromium. Which color do you think will cause the emission of an additional photon?

If you see two photons moving to the right and the electron in the chromium atom returns to energy level E_1, you have successfully simulated the workings of a laser. Congratulations! The second photon you fired by pressing GO has been "amplified" since it results in two photons moving through the ruby. It has taken advantage of the energy stored in the atom to do this.

As you conduct your experiments, you may note that the details of the laser process differ from material to material. With neon, the crucial transition is from the highest metastable level to a middle level. With chromium, the material used for this transition, the crucial transition is from a middle metastable level back to the lowest energy state.

Bohr atom: The atom consists of a nucleus surrounded by electrons orbiting it at specific radii and energy levels.

Einstein showed that light was quantized. The Danish physicist Niels Bohr proposed a model of the atom in which the energy of electrons was quantized, and could only exist at certain values. Together, these theories explained the frequencies of spectral emission and absorption lines.

Before discussing Bohr's theory, we will briefly explain some work that preceded his, and then explain his crucial hypotheses.

In 1897, the scientist J. J. Thomson showed that the "rays" often observed flowing between charged electrodes in a vacuum were streams of negatively charged particles. In other words, he discovered the electron. His discovery caused scientists to update their model of the atom. Some theorized that atoms consisted of a "mix" of negative particles and positive regions, like negatively charged chocolate chips embedded in positively charged cookie dough. (In fact, Thomson called it the *plum pudding model*, after the plums scattered throughout a pudding.)

However, in 1910 Ernest Rutherford conducted experiments that led scientists to reject the plum pudding model. He fired alpha particles at a thin gold foil, expecting them to sail through. (An alpha particle consists of two protons and two neutrons, and is positively charged.)

Rutherford discovered that although most of the particles passed through the foil with minimal deflection, a few had violent collisions. These particles were deflected at extreme angles, or even rebounded straight back.

Only a dense, positive nucleus could explain this result. Rutherford's subsequent analysis led him to conclude that this tightly packed nucleus must be surrounded by electrons that were orbiting the nucleus at relatively great distances, a model eerily similar to the solar system. In short, he developed the basis of a model that is still commonly used today to describe the atom.

During the period of these discoveries, the physicist Robert Millikan measured the magnitude of the elementary charge, which is the amount of charge of an electron or a proton. In a little over a decade, the basic model of a positively charged nucleus orbited by negatively charged electrons had been established, as had the charge of an electron.

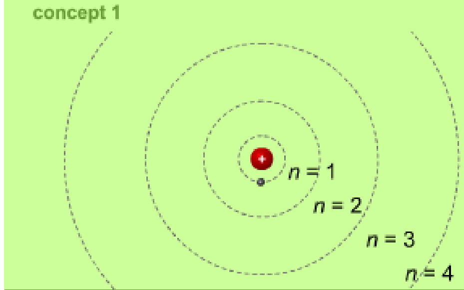
The model of the atom was radically advanced by the work of Thomson, Rutherford and Millikan. However, physicists remained puzzled by a paradox stemming from their understanding of electromagnetic theory and orbital mechanics.

Electromagnetic theory predicts that accelerating charges emit electromagnetic radiation. Electrons circling around the nucleus of an atom are constantly accelerating because they are constantly changing direction; this is similar to how the electrons in an antenna repeatedly accelerate back and forth as they oscillate over its length in simple harmonic motion. If orbital electrons were emitting radiation due to their acceleration, they would be losing energy, and they should eventually crash into the nucleus. The analogous effect is witnessed with a satellite orbiting the Earth: If it continually loses energy due to atmospheric resistance, its orbital radius decreases, and it eventually crashes.

However, since most atoms are stable (phew!), electrons are not "crashing" into nuclei, but rather are maintaining orbits of a constant radius. Bohr could not explain why the electrons acted as they did, but he formulated a theory that was consistent with what physicists were observing. He stated that electrons in atoms could only exist in certain orbits called *stationary orbits* or *stationary states*. Bohr postulated that in these states the size and energy of an electron's orbit is stable and constant. (By using the term stationary, Bohr did not mean that the electrons stood still, but rather that their orbital radii and energies remained constant.)

In Bohr's model, both the orbital radius of an electron and the total energy of its orbit are quantized. Electrons can only exist at certain distances from the nucleus that correspond to certain energy levels. (Bohr used the concept of angular momentum to determine the sizes of the orbits.) His model also led to the conclusion that an electron in such an orbit does **not** radiate energy continuously and that its orbital radius cannot gradually decay, but remains constant unless it is disturbed.

concept 1



Bohr atom
Electron moves in circular orbit around nucleus
Change in energy alters orbital size
Orbits are quantized

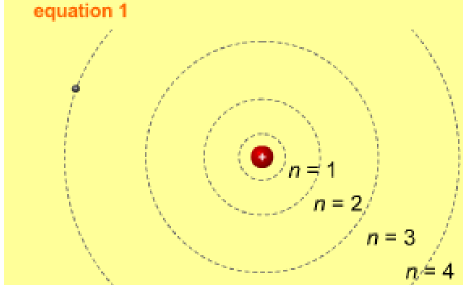
concept 2

Table of energy states of hydrogen electron

n	Energy of electron (eV)
1	-13.6
2	-3.40
3	-1.51
4	-0.85

Energy levels of hydrogen atoms

equation 1



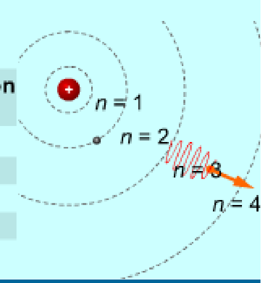
Energy levels of hydrogen atoms

$$\Delta E = E_f - E_i$$

ΔE = change in energy of atom
E = energy of an electron's orbit

example 1

n	Energy of electron (eV)
1	-13.6
2	-3.40
3	-1.51
4	-0.85



When an electron moves from N_3

In Concept 2 we show a table of energy levels for a hydrogen atom. The lowest energy level is called the *ground-state energy level*. At this level, the electron is at its closest to the nucleus, and this distance is called the *Bohr radius*. This is the smallest possible orbit of the electron. The ground-state energy of a hydrogen atom is −13.6 electron volts.

Note that the value is negative. Physicists liken this to an electron being placed in a well. It takes 13.6 eV to remove the electron from the atom so that it is free, no longer bound to the proton that is the nucleus of the hydrogen atom. The closer it is to the nucleus, the more negative its energy. This is akin to measuring the gravitational potential energy of, say, a rock at the bottom of a well. Its gravitational potential energy is stated to be negative there, and it becomes less negative as it approaches the top of the well at the surface of the Earth.

Today, electrons are not considered to be particles moving like satellites in orbits around the nucleus, and a quantum-physics *electron cloud model* of the atom has replaced the Bohr model. Nevertheless, many of Bohr's ideas have proven to be very useful, and his model greatly advanced the understanding of the atom.

How does Bohr's work relate to spectral lines? His model provided the first steps toward a quantized view of the atom. As an electron moves between specific energy levels, it emits or absorbs a quantized amount of energy in the form of a single photon of a specific frequency.

The spectral lines that result when a gas emits or absorbs energy are thus also quantized. Bohr's work provided a model on the atomic side of why this should be the case.

to N_2, what happens to its energy? Calculate the change in energy of the atom.

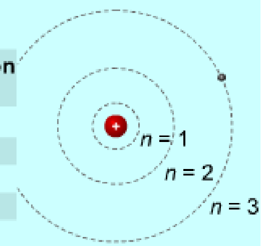
$\Delta E = E_f - E_i$

$\Delta E = (-3.40\text{ eV}) - (-1.51\text{ eV})$

$\Delta E = -1.89\text{ eV}$

example 2

n	Energy of electron (eV)
1	-13.6
2	-3.40
3	-1.51
4	-0.85



By how much does the atom's energy change when the photon strikes, moving the electron from N_1 to N_3?

$\Delta E = E_f - E_i$

$\Delta E = (-1.51\text{ eV}) - (-13.6\text{ eV})$

$\Delta E = 12.1\text{ eV}$

41.10 - Derivation: Bohr radius

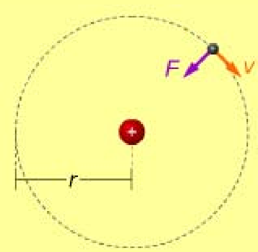
Bohr concluded that the electrons in an atom maintain orbits of certain radii. In other words, he stated that the radii (and energies) of the electrons are quantized. How did he determine that these properties are quantized, rather than continuous? (By way of contrast, classical physics allows the orbital height of a satellite above the Earth to vary continuously: Orbits of 250.0, 250.1 or 250.3 km above the planet's surface are all quite possible.)

In this section, we use classical mechanics, together with Balmer's observationally based description of a series of lines in the spectrum of hydrogen gas, to derive some properties of the electron orbits in a hydrogen atom. The variables that express the electric force on an electron and the radii of the orbits in a hydrogen atom are displayed in Equations 1 and 2 on the right; other variables are listed below.

Variables

mass of electron	$m_e = 9.11\times10^{-31}$ kg
tangential speed of electron	v
Coulomb's constant	$k = 8.99\times10^9\text{ N·m}^2/\text{C}^2$
charge of particle	q
elementary charge	$e = 1.60\times10^{-19}$ C
kinetic energy of electron	KE
potential energy of electron	PE
total energy of electron	E
energy of photon	E_p
Planck's constant	$h = 6.63\times10^{-34}$ J·s
frequency of spectral line	f
speed of light	$c = 3.00\times10^8$ m/s
wavelength of spectral line	λ
Rydberg constant	$R_H = 1.097\times10^7\text{ m}^{-1}$

equation 1



Electric force on electron

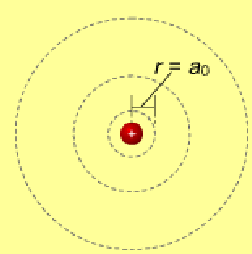
$$F = \frac{ke^2}{r^2}$$

F = force, k = Coulomb's constant

e = elementary charge

r = proton-electron distance

equation 2



Radii of orbits in a hydrogen

Strategy

1. In a hydrogen atom there is one electron and a nucleus consisting of one proton. Assume that the electron is moving in a circular orbit around the proton and state the centripetal force acting on the electron. By comparing this to the electric force exerted on the electron by the nucleus, calculated with Coulomb's law, find the kinetic energy of the electron in terms of the elementary charge and the radius of its orbit.
2. Determine the total energy of the electron in its orbit, its kinetic energy plus its potential energy. The potential energy, and hence the total energy, can also be expressed in terms of the elementary charge and the radius of the orbit.
3. Use the total energy equation that you just derived and state, in terms of the radii involved, a **theoretical** equation for the energy released in the form of a photon when the electron drops from an initial orbit of higher energy to a final orbit of lower energy. In contrast, combine Planck's and Balmer's equations to state an **empirical** equation for the energy of the emitted photon in terms of the initial and final orbital numbers.
4. Compare the two equations derived in the previous steps for the energy released as an electron changes orbits. From this comparison deduce an equation for the radii of allowed orbits, in terms of the orbital number *n* and the radius of the smallest orbit (the Bohr radius). Also, in a final step, compute the Bohr radius.

atom

$$r_n = n^2 a_0, \quad n = 1, 2, 3, \ldots$$

r_n = radius of nth orbit
n = orbital number
a_0 = Bohr radius (smallest orbit)
Radius $a_0 = 0.529\times10^{-10}$ m

Physics principles and equations

The formula for centripetal force in uniform circular motion

$$F = \frac{mv^2}{r}$$

Coulomb's law, stated for an electron and a proton

$$F = k\frac{|q_1||q_2|}{r^2} = k\frac{e^2}{r^2}$$

The definition of kinetic energy

$$KE = \frac{1}{2}mv^2$$

The equation for the PE of a system comprising two opposite charges, an electron and a proton, separated by a distance r, with the PE defined to be equal to 0 at infinity

$$PE = \frac{kq_1q_2}{r} = -\frac{ke^2}{r}$$

The total energy of an orbiting electron is the sum of its kinetic and potential energies,

$$E = KE + PE$$

Planck's equation for the energy of a photon in terms of its frequency or wavelength

$$E_p = hf = hc/\lambda$$

The Balmer equation

$$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$$

Step-by-step solution

First, we find the kinetic energy of the electron in terms of its charge and orbital radius.

Step	Reason
1. $F=\frac{m_e v^2}{r}$	centripetal force
2. $F=\frac{k\|q_1\|\|q_2\|}{r^2}$	Coulomb's law
3. $F=\frac{ke^2}{r^2}$	e is elementary charge
4. $\frac{m_e v^2}{r}=\frac{ke^2}{r^2}$	combine equations 1 and 3
5. $\frac{m_e v^2}{2}=\frac{ke^2}{2r}$	multiply by $r/2$
6. $KE=\frac{ke^2}{2r}$	definition of KE and equation 5

In order to calculate the total energy of the electron we must also calculate the potential energy of the proton-electron pair in the hydrogen atom.

Step	Reason
7. $PE=\frac{kq_1q_2}{r}=-\frac{ke^2}{r}$	potential energy of system
8. $E=KE+PE$	total energy
9. $E=\frac{ke^2}{2r}-\frac{ke^2}{r}=-\frac{ke^2}{2r}$	substitute equations 6 and 7 into equation 8

We have found the total energy of an electron orbiting at a particular radius. Bohr's first postulate was that energy is quantized for an electron in a hydrogen atom, so there are only certain radii at which the electron can be orbiting. Bohr's second postulate was that in a quantum jump between orbits the electron emits (or absorbs) radiation in the form of a photon. We use these postulates to create an expression for the energy released when an electron drops from a higher energy level to a lower energy level. Then we compare this to the photon energy implied by the Balmer equation, which is based on experimental data.

Step	Reason
10. $E=\left\|-\frac{ke^2}{2}\left(\frac{1}{r_f}-\frac{1}{r_i}\right)\right\|=\frac{ke^2}{2}\left(\frac{1}{r_f}-\frac{1}{r_i}\right)$	energy emitted as electron changes radius
11. $E_p=hf=\frac{hc}{\lambda}$	energy of photon
12. $\frac{1}{\lambda}=R_H\left(\frac{1}{2^2}-\frac{1}{n^2}\right)$	Balmer equation
13. $E_p=hcR_H\left(\frac{1}{2^2}-\frac{1}{n^2}\right)$	substitute equation 12 into equation 11

We will now compare the "theoretical" energy-difference equation (step 10) to the "observed" photon energy equation (step 13). The equations look similar and we will exploit this to obtain an expression for the radii of the electron orbits. In step 10 there are two variables r_i and r_f, the radii of the initial and final orbits, while there is just one variable n in step 13, describing the initial energy state. That is, equation 13 is a specific case of the more general equation 10.

In fact it is known that the Balmer series describes an electron dropping from a higher energy level to the 2nd level. (There are other hydrogen spectral series, lying outside the range of visible light, for electrons dropping to levels other than 2.)

Step	Reason
14. $\frac{ke^2}{2}\left(\frac{1}{r_\mathrm{f}}-\frac{1}{r_\mathrm{i}}\right)=hcR_\mathrm{H}\left(\frac{1}{2^2}-\frac{1}{n^2}\right)$	combine equations 10 and 13
15. $\frac{ke^2}{2}\left(-\frac{1}{r_\mathrm{n}}\right)=hcR_\mathrm{H}\left(-\frac{1}{n^2}\right)$	equate the "initial" terms in step 14
16. $r_\mathrm{n}=n^2\left(\frac{ke^2}{2hcR_\mathrm{H}}\right)=n^2a_0$	simplify and solve for r_n
17. $r_1=1^2a_0=\frac{ke^2}{2hcR_\mathrm{H}}$	Bohr radius
18. $a_0=\frac{\frac{1}{2}(8.99\times10^{9})(1.60\times10^{-19})^2}{(6.63\times10^{-34})(3.00\times10^{8})(1.097\times10^{7})}$ $a_0=0.527\times10^{-10}\mathrm{m}$	evaluate Bohr radius

The value for the Bohr radius derived above is very close to the experimental value, which is currently accepted as 0.529 177×10^{-10} m. There is a small error in our calculation because we assume that the proton has, in essence, an infinite mass and does not wobble as the electron orbits around it. Though the proton's mass is not infinite, it is very large compared to the electron's mass (roughly 2000 times as large), and this is a good approximation.

Bohr did not actually perform the derivation as you see it here (see the following derivation), but he is reported to have said, "As soon as the Balmer formula was shown to me, the whole thing became clear in my mind."

41.11 - Energy levels, photons and spectral lines

Hot, excited metal atoms emit light at characteristic frequencies. For example, fireworks packed with copper salts radiate blue light when they explode.

The Bohr model, combined with Einstein's and Planck's work, explained the discrete spectral lines that physicists were observing.

Why does excited hydrogen gas only emit light at certain frequencies? When it is excited, say by heat, or by an electric current, its atoms absorb energy. When a hydrogen atom absorbs energy, its electron jumps from one orbit to another, from a lower energy level to a higher one. The change in energy is quantized because electrons can only exist at the specific energy levels prescribed in Bohr's model.

When the atom loses energy, it does so by releasing a single photon. The energy of the photon corresponds to the amount of energy the electron loses as it returns to a lower-energy orbit.

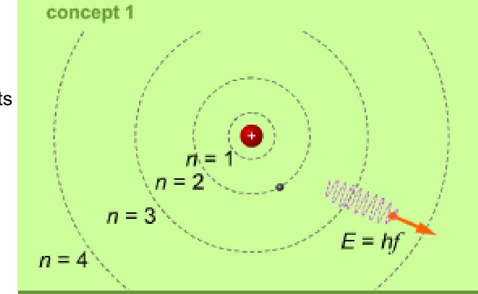
The energy of the photons emitted by excited gas atoms must be quantized because the electron energy levels are quantized. The frequency of a photon is proportional to its energy, $f = E/h$. We perceive a given frequency (or wavelength) of light as a specific color.

For example, consider the red-colored line having wavelength 656 nm and frequency 4.57×10^{14} Hz in the emission spectrum of hydrogen. A photon of this color must have energy 3.03×10^{-19} J, or 1.89 eV.

It is possible to calculate the orbital change of the electron in a hydrogen atom that creates light of this frequency. An electron in the $n = 3$ orbit has −1.51 eV of energy. An electron in the $n = 2$ orbit has −3.40 eV of energy. An electron that drops from N_3 to N_2 gives up 1.89 eV of energy: exactly the energy of the red-color photon. A mystery solved! Some Nobels won!

In sum, in the first decades of the 20th century, scientists had already discovered the physics that Andy Grove and his peers would need in order to create semiconductors and lasers. In fact, the first patents for basic semiconductor transistors were granted in Germany and the United States in the 1920s and 1930s. The practical manufacture of these devices required progress in the material sciences, but transistors were

concept 1

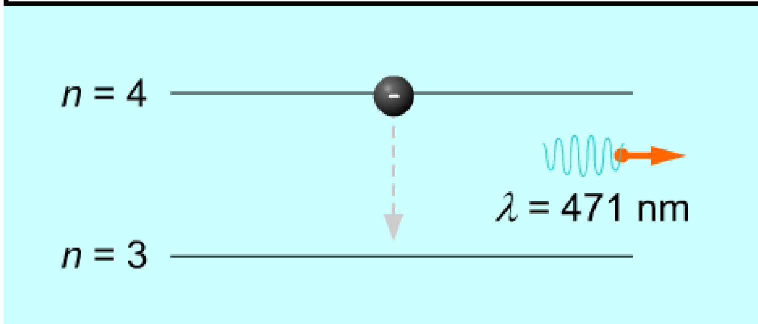


Bohr, photons and spectral lines

Orbits (and energy levels) are quantized
· Only certain energy changes possible
Photons of light are quantized
· Energy = Planck's constant times frequency
· Observed spectral lines match energy differences

manufactured in the late 1940s, and the first modern-day field effect device, a type of transistor, was proposed in 1952 and built in 1953.

41.12 - Interactive checkpoint: photons and electron energy levels



A helium atom emits light of wavelength 471 nm when an electron makes a transition from the $n = 4$ state to the $n = 3$ state. If the energy of the $n = 3$ state is -6.04 eV, what is the energy of the $n = 4$ state?

Answer:

$E =$ ______ eV

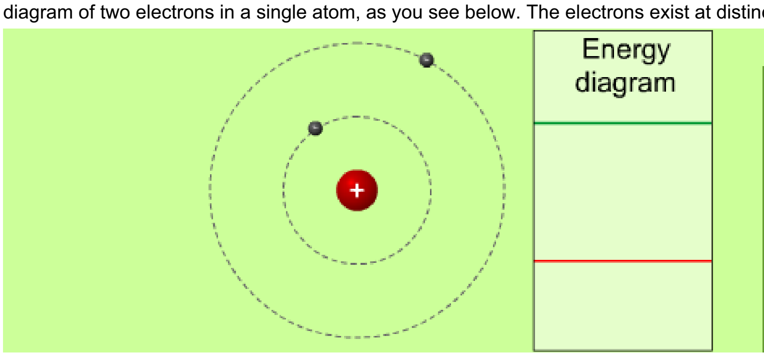
41.13 - Conduction in solids

This chapter started off by noting that the development of semiconductor-based devices, such as transistors, has been one of the most important technological advances of the past 50 years. We also promised that understanding their functioning would require only a basic understanding of some principles of quantum physics. With your introduction to quantized energy levels and photons essentially complete for the purposes of this chapter, we can now start the discussion of semiconductors, and shortly, lasers.

To explain semiconductors, we need to discuss the conduction of electrical currents in solids. In terms of conducting electricity, semiconductors lie between the extreme cases of conductors and insulators (hence the “semi” in their name). The ability of engineers to influence how readily semiconductors conduct a current is a key to their utility.

To review some terminology: Some substances (like copper or aluminum) are considered conductors; current flows relatively easily in a conductor. Others (like silicon laced with impurities) are considered semiconductors. Others, like silicon dioxide, are insulators, where it is very difficult to cause a current to flow.

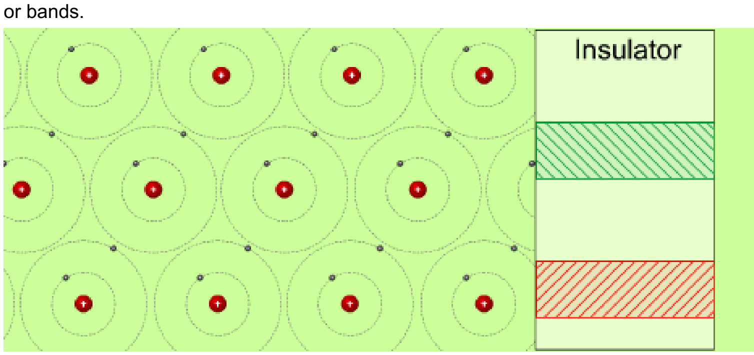
In this section, we discuss why current flows more − or less − easily in conductors, semiconductors and insulators. We start with an energy diagram of two electrons in a single atom, as you see below. The electrons exist at distinct energy levels.



concept 1

Energy levels of a single atom

To understand the relative ease or difficulty of causing a current to flow, we need the concept of *energy bands*. Below we show what happens when many atoms are brought together, say atoms that have crystallized into a solid. When atoms are brought very close together, their energy levels interact and merge into a set of broader energy bands. Instead of the electrons of an individual atom being restricted to particular energy levels that are so sharp and distinct they are represented with lines, the electrons now can exist with their energies in ranges of values, or bands.

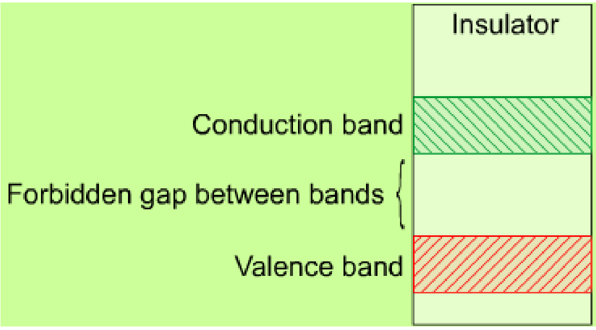


concept 2

Energy bands in a solid
Presence of many atoms broadens energy levels into bands

In fact, we can get more specific about these bands. We will use the diagram below to do so. As you can see, the lower energy band is the *valence band*. If you have studied chemistry, you may recall that valence electrons are those most likely to participate in chemical bonds. Although atomic electrons in the valence band can be shared with neighboring atoms in chemical reactions, they are too strongly bound to the nucleus of their own atoms to be able to flow freely in an electric current. Electrons at lower energy levels are even more closely tied to the nucleus, cannot flow in current, and are irrelevant to our discussion.

The highest-energy electrons are in the *conduction band*. It is electrons in this band that can flow in a current. They are crucial to our story: Electrons in the valence band cannot flow in a current; electrons in the conduction band can.



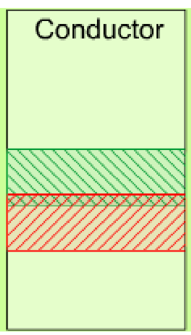
concept 3

Two outer bands of insulator
Valence band
Conduction band
Large energy gap between bands
· "Forbidden gap"

It takes energy to move from one band to another, just as it takes energy for an electron to move from one energy level to another in an isolated atom. When energy is added to an atom, its electrons can be promoted from the valence band to the conduction band, where they are free to move and become part of a current. How much energy it takes to accomplish this determines whether the material is classified as a conductor, a semiconductor or an insulator.

Let's now discuss this concept with specific materials. The diagram above shows the *energy band diagram* of an insulator, say silicon dioxide, for the valence and conduction bands. As in the prior illustration, the valence electrons occupy states in the lower energy band shown in the diagram. The conduction band is the upper band. Between these two bands lies the *forbidden gap,* or *band gap*. This is the Mojave Desert of electrons. They are not allowed to exist there. Electrons can only exist with energies in ranges like the valence or conduction bands.

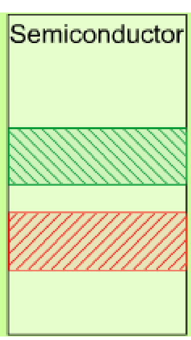
As you see, the forbidden gap is comparatively large in an insulator. It takes a relatively large amount of energy (8 eV for silicon dioxide, if you like specifics) to cause an electron to move from the valence band to the conduction band. With few electrons in the conduction band, it is extremely difficult to produce a current.



concept 4

Conductor
Energy bands
· Bands overlap

Now we consider the other extreme: a conductor. Its energy diagram is shown above. Note how the energy bands overlap; in a conductor, this overlapping region is only partially filled with electrons. It is "easy" for an electron to move from the valence to the conduction band when a potential difference is introduced. This means a conductor like copper can provide a ready supply of conduction electrons, ready to rumble when the slightest electric field from a source like a battery is applied.

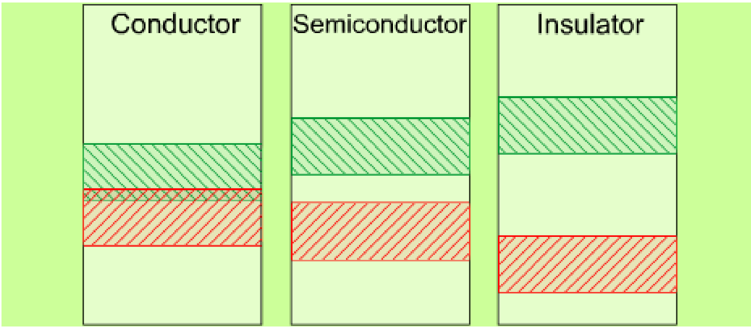


concept 5

Semiconductor
Energy bands
· Medium gap between bands

Above, you see the energy level diagram for the third type of material, a semiconductor. (We use silicon; purists may rightly complain that pure silicon is not truly a semiconductor, but we hope they will let us slide for a moment.) Its properties lie between those of insulators and conductors. As the diagram suggests, it takes less energy than with an insulator for an electron to move from the valence to the conducting band. At room temperature, it takes 1.12 eV.

To contrast the differences, the diagram below shows energy band diagrams for the three types of material side by side by side. You can see how the energy gaps differ by the type of material.



concept 6

Three types of material
Contrast between energy gaps

41.14 - Mobile electrons and holes

A hole: is caused by the departure of an electron and is positive.

In a physics topic like direct current circuits, an electric current is described as a flow of electrons, since it is moving electrons that are the charge carriers in a copper wire. However, semiconductor engineers think about the flow of current slightly differently. We will use a silicon atom or two to explain.

Here is a traditional discussion of *covalent bonds* as taught in first-year chemistry: A single silicon atom has four electrons in its valence band, but it “wants” to have eight there. When it is part of a solid piece of silicon, it forms covalent bonds with neighboring atoms. It “shares” electrons with neighboring silicon atoms, and by sharing, it fills its valence band with eight electrons. You see a symmetrical, “satisfied” silicon atom in the diagram of Concept 1.

Now let us consider what happens when an electron makes a jump from the valence band to the conduction band. It might do so as the silicon increases in temperature and the internal energy of the material increases.

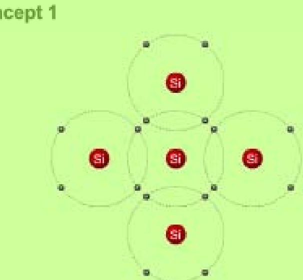
When an electron makes this jump, a silicon atom has lost its valence electron. The valence band now has an opening, called a hole. The diagram in Concept 2 shows the “missing” electron as a hole. A hole is positive since it is caused by the departure of an electron. The number of protons in the region now exceeds the number of electrons by one.

Holes are crucial in semiconductors. Why? Because they provide a place for electrons to flow to. They provide natural “landing spots” for mobile electrons.

However, there is more to it than that. When considering semiconductors, the flow of electrons is crucial, as electrons constitute the electrical current in a conductor. However, equally important and real in the eyes of semiconductor engineers is the flow of holes. In the animation in Concept 3, we illustrate some moving electrons and the resultant motion of a hole. Refresh your browser screen if you did not see this animation yet and wish to do so.

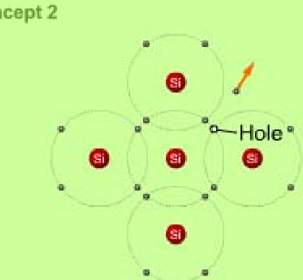
You may consider hole movement as being akin to a bubble moving through a fluid. A bubble is the absence of fluid. The motion of the bubble can be described more concretely as a movement of the fluid around it. When the bubble moves one way, there is a flow of fluid in the other direction. However, the “movement” of the bubble is much more noticeable than the motion of the fluid itself. Semiconductor engineers and physicists treat holes as though they were particles as real as electrons.

concept 1



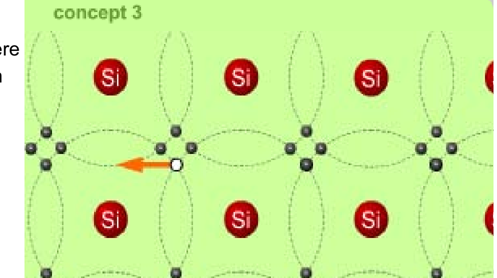
Silicon in equilibrium
Valence band is filled

concept 2



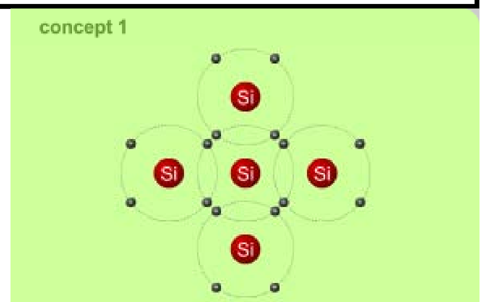
Silicon with an electron in the conduction band
Valence band is “missing” an electron
A hole is created
· Holes are positive

concept 3



Holes resemble bubbles
Holes move

Doping: Increasing the availability of conducting electrons or holes by adding impurities to a material like silicon.



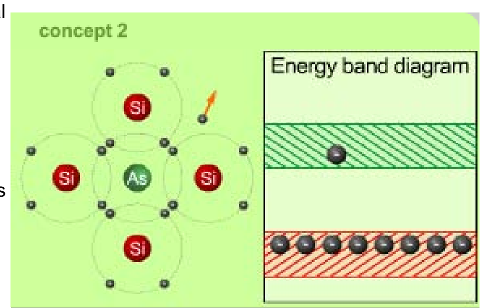
Pure silicon

Commercial semiconductors are *doped*. This is thoroughly frowned upon in the Olympics, but quite a desirable thing to do in a semiconductor foundry. Adding impurities to a substance like silicon allows engineers to tailor its electrical properties and makes it more useful in building devices like diodes and transistors.

We start with the diagram in Concept 1 showing silicon atoms in an ideal or equilibrium state, perhaps at a temperature near absolute zero. The atoms fill their valence bands by sharing electrons.

Now let's consider the same pure silicon, but at room temperature. The average thermal energy of the atoms has increased. This increase in energy means some electrons will have enough energy to spontaneously make the energy jump from the valence to the conducting band.

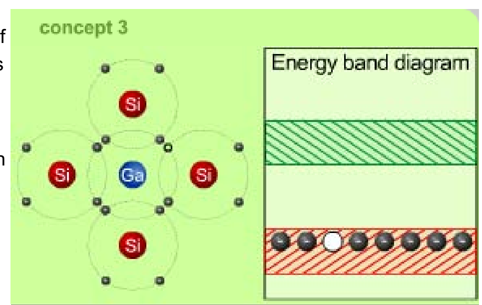
In fact, in absolute terms, a large number of electrons will make the jump. A cubic centimeter of silicon contains 5×10^{22} atoms, and 2×10^{23} valence electrons. At room temperature (293 K), you will find about 1×10^{10} electrons in the conducting band in this volume of silicon. Since they have left the valence band, you will find an identical number of holes. On the one hand, this is a vast number ($10{,}000{,}000{,}000$ electrons and an equal number of holes). On the other hand, it is extremely small compared to the total number of electrons: About one electron out of every 10^{13} valence electrons has become a conducting electron.



Doping
Adding a different material to silicon
n-type: inserts extra electron
"Extra" electron readily moves to conduction band

One out of 10^{13} is a small fraction, less than semiconductor engineers desire. To increase the number of conducting electrons and holes, the silicon is doped. In doping, impurities are added to the silicon to increase its ability to conduct current. Arsenic and phosphorus are common elements that are added to increase the number of available conducting electrons, while hole-increasing elements include gallium and boron.

Arsenic and phosphorus atoms both have five valence electrons, one more than silicon has. When an arsenic atom takes the place of a silicon atom in the atomic lattice, four of its valence electrons fit easily into the covalent bonds and in essence become members of the valence band of the adjacent silicon atoms. The fifth electron, however, enters the conduction band since there is no room for it in the valence band.



Creating holes with doping
p-type: an electron "short"
Provides destination for mobile electrons

In Concept 2, you see an arsenic (As) atom and its "fifth" electron. When it is doped with an element like arsenic, the semiconductor is called an *n-type* semiconductor. The "*n*" stands for negative, since the charge carriers supplied by the dopant are negative. The "extra" electrons are called *donor* electrons.

To facilitate the flow of current, an engineer may want a material with holes: places for those electrons to flow to. She would use an element like gallium or boron whose atoms have just three valence electrons. When such an element replaces silicon in the atomic lattice, it is one electron short. The result is a hole, as shown in Concept 3.

When it is doped with an element like gallium (Ga), the semiconductor is called a *p-type* semiconductor. The "*p*" stands for positive, representing the fact that holes act like positive charges when they move.

Arsenic is used to supply carrier electrons and gallium is used to supply holes. In quantitative terms, the band gap separating the "fifth" arsenic valence electron from the conduction band is about 1/20th the size of the band gap for the silicon valence electrons.

In sum, an *n*-type semiconductor has a set of electrons ready to move, while in contrast a *p*-type semiconductor has a place ready for valence electrons to go, freeing up holes that can then move. When an external electric field is applied to a doped semiconductor, current will flow much more readily than in a pure semiconductor.

41.16 - p-n junction

p-n junction: *p*-type and *n*-type semiconducting materials placed adjacent to one another. This type of junction is the basis of devices like diodes and transistors.

Diode: A component that readily allows the flow of current in one direction,

and is highly resistant to current in the other.

In this section, we discuss what happens when *p*- and *n*-type materials are placed in contact with one another. One result is a useful device, the diode.

Consider the *p*-*n* junction shown in Concept 1. A junction refers to the region or a device where the two types of semiconducting material are touching. Remember that the *n* section has excess electrons that can flow fairly readily, while the *p* section has excess holes that could accept those electrons.

When these two types of material are placed next to one another, some holes flow from the *p* to the *n* material, and some electrons flow from the *n* to the *p* material.

This is called a *diffusion current*. The electrons diffuse from the *n*-type material, where there is a higher concentration of them, to the *p*-type, where they are relatively scarce. The holes move in the opposite direction, from the *p*-region where they are abundant to the *n*-region where they are scarce. They diffuse, just as perfume molecules diffuse across a room.

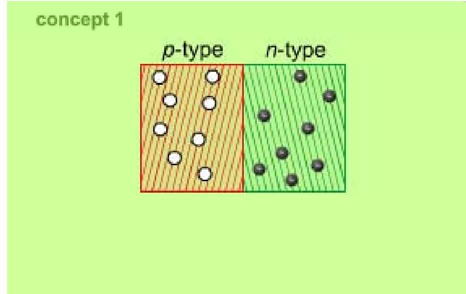
The *p*-*n* junction is the essential element of a diode. When it is connected in a circuit in the fashion shown in Concept 2, the net flow of current is relatively great at low voltages. This is called a *forward-bias connection*.

Why does the current flow easily? Electrons flow from the *n*-region of the semiconductor rather readily to fill holes on the *p* side. The negative terminal of the battery acts to replenish the supply of electrons in the *n*-region and the positive terminal replenishes the holes in the *p*-region.

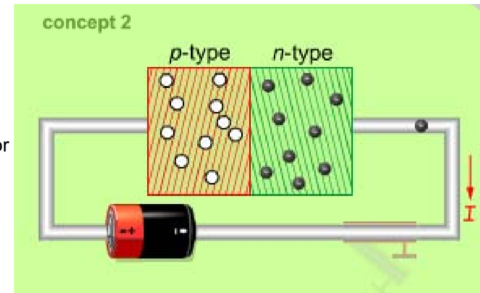
Now we reverse the orientation of the battery, as shown in Concept 3. The diode will block nearly all the current for low applied voltage: It acts as a resistor of great resistance. This is called a *back-bias connection*.

Why does this orientation prove so resistant to current? As Concept 3 shows, the battery causes a significant *depletion zone*. Holes in the *p*-region are moved away from the junction as they move toward the negative terminal of the battery, and free electrons in the *n*-region move away from the junction toward the positive terminal. The region around the junction on both sides loses its mobile charge carriers; it becomes depleted. The battery can “pull” harder and harder, but in effect, all it does is expand the depletion zone, instead of causing a continuing current.

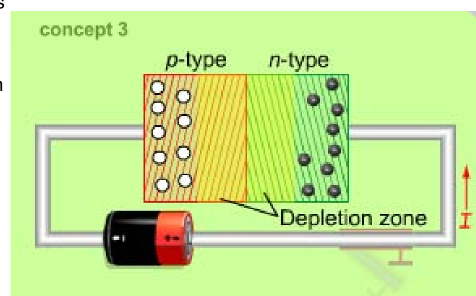
The two graphs on the right show current versus voltage curves for forward bias and reverse bias connections. In the forward bias case, the current increases with potential difference. The diode acts roughly like a resistor. In the reverse bias case, the diode acts almost like a break in the circuit, and even relatively large potential differences cause negligible currents.



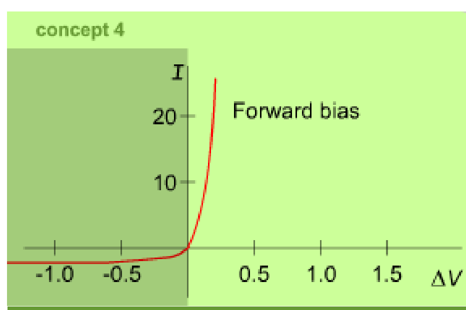
p-n junction
p-type material has excess holes
n-type material has excess electrons



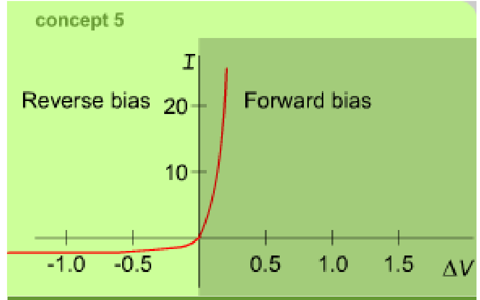
Diode
Allows current to flow in one direction



Battery reversed
Creates depletion zone
Prevents flow of current

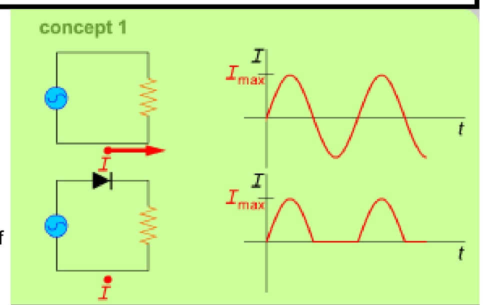


Forward bias
Current increases with ΔV



Reverse bias
Current does NOT increase with $|\Delta V|$

41.17 - Physics at work: junction rectifiers



Junction rectifier
Turns AC into DC

Junction rectifier: A device that allows current to flow in one direction but not the other.

What might be an application of a diode? One practical application is to use the diode as a junction rectifier.

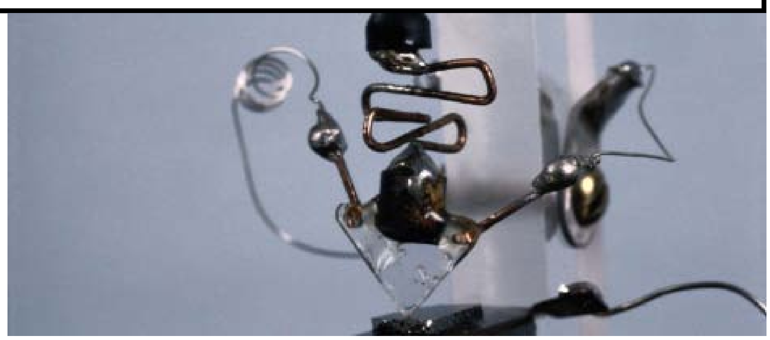
An alternating current (AC) generator produces a potential difference (or voltage, or emf if you like) that oscillates over time, oriented in one direction for half of each cycle, and switching to the other direction for the other half. The top graph in Concept 1 shows the graph of the current over time in a simple circuit that contains a generator and a resistor.

Say you want to charge the battery of a laptop by plugging it into an electrical outlet. When you plug the laptop into the outlet, you are in essence connecting it to an AC generator. However, like many devices that run on batteries, the laptop battery itself requires direct current (DC), a steady current that flows in just one direction, in order to charge up.

If you could just use half of each cycle produced by the AC generator, say the part consisting of a peak, you would be on your way to converting AC to DC. You could accomplish this with a diode, which only allows current to flow in one direction.

You see this in the lower graph of Concept 1, where we show the current caused by placing a diode in a simple circuit with an AC generator and a resistor. All the troughs are blocked, and only the peaks remain. Other components (such as capacitors) can be placed in the circuit to smooth out the peaks in the emf and produce a nearly constant voltage.

41.18 - Physics at work: MOSFET transistors



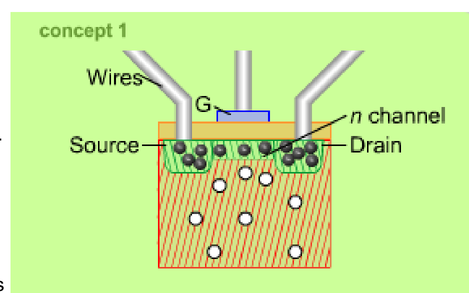
The first transistor, built at Bell Labs in 1947. The semiconductor substrate is a germanium crystal. The three electrical leads are the source, gate and drain.

Transistor: A three-terminal semiconductor device that forms the basis of random access memory and microprocessors.

The physicist William Shockley proposed the first modern-day transistors in 1952, although earlier scientists had devised conceptual prototypes. In a circuit, transistors can be controlled so that they work either as strong resistors or effective conductors, which is what makes them important in electronics applications, such as electronic on/off switches (in computers), or as part of signal amplifiers (in your stereo).

The transistors discussed in this section are called *field effect transistors*, since the conductivity (or resistance if you like) of such transistors is regulated by an electric field. The first transistors, developed earlier, were of a type called *junction transistors* or *junction field effect transistors*, but *MOSFET* transistors dominate many applications today.

MOS stands for metal-oxide-semiconductor and FET stands for field-effect transistor. MOSFET transistors are crucial in microelectronics, forming the basis of random access memory (*RAM*) and of *charge-coupled devices* (CCDs). A *CCD* is used to record images in digital cameras and digital video cameras.



Structure of a MOSFET transistor
Substrate is lightly doped *p*-type

Enough acronyms: There are many more! (Other acronyms were suggested for

MOSFET, including MISFET…) Let's move to the design of an *n channel* MOSFET transistor. The basic design is shown in Concept 1.

Silicon is lightly doped to form a *p*-type semiconductor, which is the bottom layer you see in the diagrams on the right. This layer, the *substrate*, is deliberately very lightly doped so that it will be a poor conductor, but the holes it contains are necessary for a reason you will soon learn. Then two much more heavily doped *islands* of *n*-type semiconductor, which are shown in green in the diagram, are formed on top of the substrate. One island is the *source* S and the other is the *drain* D.

A thin channel of *n*-type material, called appropriately enough, the *n channel*, connects the two islands. A layer of insulating material (such as silicon di**o**xide, whence the "O" in MOSFET), mere nanometers thick, is deposited on top and penetrated by the two **m**etal leads (whence the M) shown in the diagram. On top of the insulating material lies a third deposit of metal, called the *gate* G. Because of the insulating layer, no charge can flow from the gate to the rest of the transistor.

This is a reasonably complex configuration. It is all the more impressive when you consider that semiconductor manufacturers are now building transistors where the gates measure 65 nanometers across, and that a single microprocessor chip can contain 500 million or more transistors. (It is a safe bet that humans have manufactured more transistors than any other device.)

The diagrams on the right show a transistor as part of a circuit. There is always a potential difference between the source and the drain. The key to how the transistor functions resides in whether there is a potential difference between the source and the gate. We will call this difference the gate voltage.

Let's consider what happens when there is no potential difference at the gate: In this case it has no effect on the other parts of the device. The illustration in Concept 2 shows this state. Electrons flow from the negative lead to the source island, then across the *n* channel to the drain island and the positive lead, because of the potential difference between the source and the drain. In short, a current flows through the transistor when the gate voltage is zero.

Now we assume that there is a negative gate voltage. This is illustrated in Concept 3. (A signal from another circuit might "turn on" this potential difference.) When it is turned on, we show the gate as negatively charged. Electrons repel each other, so the field caused by the electrons in the gate drives the *n* channel's electrons into the *p*-doped substrate where they find holes to occupy.

The channel is said to be *depleted*: The *n* channel has less *n*, electrons. Another way to put it is that the channel becomes narrower. This increases its resistance, and with a strong enough electric field from the gate, no electrons can flow from the source to the drain at all. This is where the "field effect" in MOSFET comes into play. (The substrate is too lightly doped and too poor a native conductor to allow current to flow there.) The entire process can be likened to stepping on a flexible garden hose to cut off the flow of water.

With a variation of the electric field of the gate, the transistor can be turned from ON to OFF, from allowing current to flow to preventing it. This simple idea, enabling circuits to be set to "ON" or "OFF", to represent "1" or "0", "true" or "false", underlies the design of computer memories and microprocessors.

After all this, you may think: It just turns on and off. Indeed. Just repeat that 500 million times or so and you have a microprocessor! The sophistication of the manufacturing process allows many transistors to be packed into a very small region.

The transistor also can function as part of an *amplifier*. Imagine that the potential difference between the source and the drain is very large: This is the "power" part of the amplifier. We show this in Concept 4.

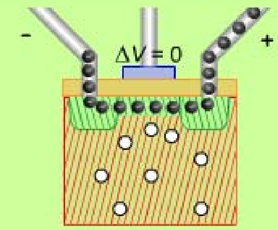
When there is no sound, the microphone creates a negative gate voltage, depleting the *n* channel and preventing current from flowing through the transistor. Although the voltage from the "large emf" source is much greater than the voltage of the microphone, the microphone is preventing any current from flowing.

Now imagine that a singer starts her song. As the sound of her voice becomes louder, the microphone creates a smaller negative gate voltage, or even a positive one, and current is allowed to flow from the drain to the source, unleashing the power of the large emf and driving the loudspeaker. The transistor functions as a variable resistor that is controlled by the microphone signal.

The strong current flowing through the channel from the source to the drain is in perfect synchronization with the amount of charge on the gate, which depends on the voltage applied by the microphone. The large signal mimics the small one.

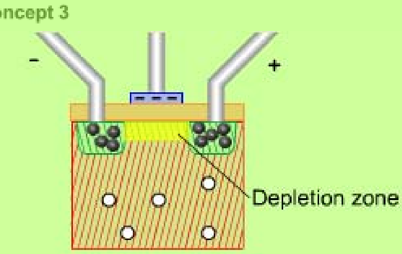
semiconductor
Two islands of *n*-type material sit atop substrate
· One island is source S, one is drain D
n channel connects islands
Insulator covers substrate, islands
Metal gate G is above insulator

concept 2



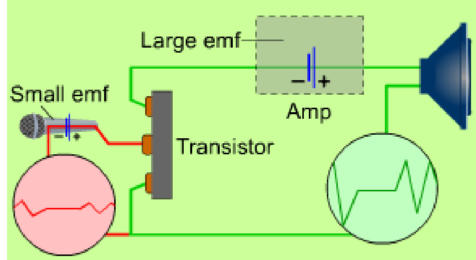
How a transistor functions (ON)
When gate voltage = 0 with
Potential difference between source and drain
· Electrons flow from source to drain

concept 3



How a transistor functions (OFF)
When gate has negative voltage
Electrons driven out of *n* channel
· *n* channel is depleted of electrons
Current cannot flow

concept 4



A transistor at work
Source/drain emf is large
Gate emf is small (signal)
Signal determines output of amplifier

Photovoltaic effect: Electron flow caused by photons.

Banks of photovoltaic cells populate a "solar energy farm" that can produce hundreds of kilowatts of electric power.

Solar cells are semiconductor devices that absorb light and convert its energy into electricity. Since sunlight is free and the operation of the cells has no environmentally hazardous side effects, there is an obvious appeal to their use. The challenge for researchers and manufacturers is to lower their cost and to raise their efficiency so that they may effectively compete with oil, coal and other traditional energy sources.

How do solar cells work? In Concept 1 we show a typical *n*-on-*p* junction solar cell, the most common type of cell. A wafer of *p*-type silicon has an element such as phosphorous diffused into its upper surface. This results in an *n*-type material, one with mobile electrons, being located above a *p*-type material, whose charge carriers are mobile holes. In short, this is a *p*-*n* junction, or diode.

With solar cells, the concept of diffusion current is important. Near the junction, mobile electrons from the *n*-region diffuse into the *p*-region, leaving positively charged donor ions (holes) in their wake. These holes remain there and form a positive region.

Mobile holes also diffuse from the *p*-region in the other direction, into the *n*-region. This flow of charge, a *diffusion current*, occurs for the same reason perfume diffuses across a room, from where the concentration is higher to where the concentration is lower. The result is a depletion zone near the junction. That is, the *n*-region is relatively depleted of mobile electrons and the *p*-region is depleted of the same number of holes.

The effects of this diffusion are shown in Concept 2. After enough charges move, equilibrium is reached: The result in the depletion zone is a built-in electric field that points from the *n*-region toward the *p*-region. This field opposes any further motion of charge, and the diffusion current quickly stops.

In an earlier section, we discussed how placing a *p*-*n* junction in a circuit with a battery could create a depletion zone. The point here is that a depletion zone also forms spontaneously when the materials are simply placed adjacent to one another.

Now let's consider what happens when a photon strikes the solar cell. The only photons relevant to our story are those that are energetic enough to promote an electron from the valence to the conduction band. When such a photon strikes a valence electron in the semiconductor, it increases the energy of the electron, promoting it to the higher band.

The electron will flow toward the *n*-type semiconductor, since that side of the depletion zone is positively charged, and the hole will flow to the negatively charged *p*-type semiconductor. To put it another way, the electric field caused by the diffusion current "pushes and pulls" the holes and electrons freed by photons striking the semiconducting material. Press the refresh button in your browser to see this occur in Concept 3.

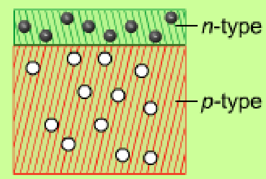
Since many photons will strike the material, many electrons and holes will flow: A current is born. By placing metallic contacts on either side of the junction, this current can be used to power a load. In Concept 4, you see the solar cell in a circuit powering a home. There is a flow of negatively charged electrons out of the *n*-region into the circuit, and a flow of positively charged holes out of the *p*-region into the circuit.

The price of photovoltaic cells has been steadily decreasing over the past 30 years. However, the electrical power produced by such cells still costs more than power from fossil fuels (coal and oil), or wind-generated power. To cite some approximate numbers (since energy prices fluctuate), a kilowatt-hour produced by burning a fossil fuel costs from 3.5 to 4.5 cents. Wind power costs just 4.5 to 5.5 cents per kilowatt-hour, although there are issues with it: What do you do when the wind is not blowing?

In contrast, the cost of solar power is approximately 25 to 45 cents per kilowatt-hour. Solar power costs more money, although it could be cost effective for supplying power to remote locations, since power lines would not have to be run from distant power plants.

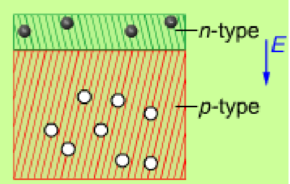
Analysts in the environmental community do raise issues about the "true" costs of different energy sources, such as the costs of environmental side effects, including health issues, pollution and global warming. Fair enough! With those costs factored in,

concept 1



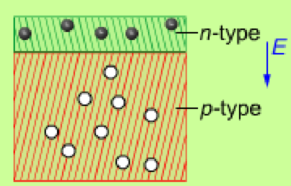
Structure of photovoltaic cell
p-substrate with *n*-type material diffused into surface
· A *p*-*n* junction

concept 2



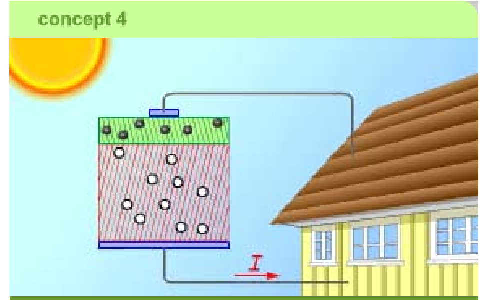
When materials placed together
Electrons spontaneously diffuse, *n* to *p*
Holes spontaneously diffuse, *p* to *n*

concept 3



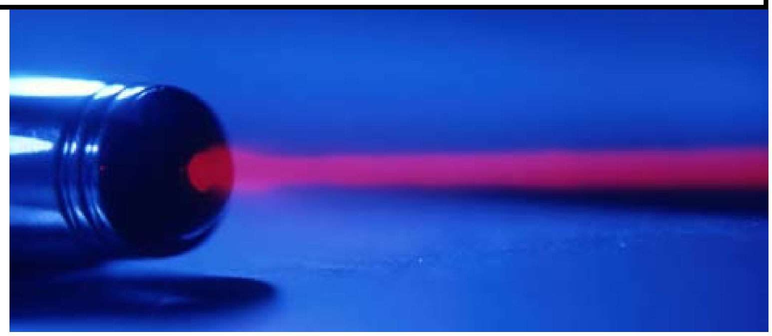
When a photon strikes
A mobile electron/hole pair is formed
The electron moves to *n*-region
The hole moves to *p*-region

different conclusions can be drawn about which power source really is the cheapest. However, the numbers above reflect the costs that companies and consumers pay in the short term for power from different sources, and they help to explain the current dominance of fossil-fuel energy.



Current flows
As photons cause electron/hole pairs

41.20 - Lasers



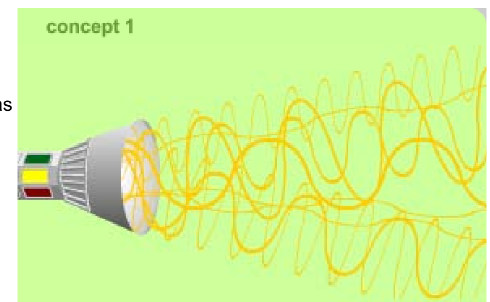
Laser emitting coherent red light.

Laser: Light Amplification by the Stimulated Emission of Radiation.

Semiconductor transistors rely on physics that was pioneered in the early 20th century, and so do lasers.

Lasers amplify electromagnetic radiation. They can be used to amplify radiation at frequencies ranging from infrared to x-rays. To simplify the discussion, we will simply say that they amplify light, not worrying about whether or not the radiation is visible to humans.

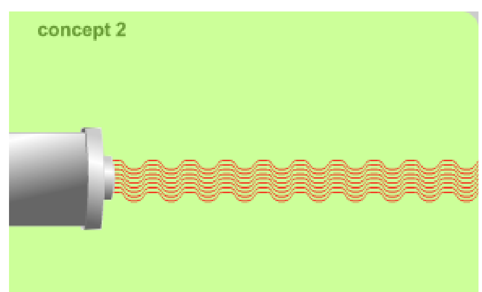
The light that emanates from a laser has three crucial properties. It is (1) coherent, (2) monochromatic, having one frequency, and (3) highly directional. *Coherent* means that the emitted waves are all in phase with one another: The light can be considered as a single wave. In contrast, the light waves that emanate from a light bulb are out of phase, or incoherent. Such light is often described as consisting of a collection of finite *wave trains*, and the trains are not synchronized.



Light from flashlight
Incoherent, multi-colored, divergent

In Concepts 1 and 2, you see light emanating from a flashlight and from a laser. The light from the flashlight consists of many wave trains, which as you can see have different wavelengths and frequencies, and are traveling in different directions. The wave trains from the flashlight are not in phase: The locations of peaks and troughs vary by wave train. The contrast with laser light is clear: its waves are in phase, have one frequency, and travel in a single direction.

A working laser has three essential parts: a laser medium, a pumping process, and a feedback mechanism.



Laser light
Coherent, monochromatic, highly directional

The *laser medium* can be manufactured from a wide variety of materials. The first operational laser used a ruby crystal. Today, the laser medium can be a gas (such as helium-neon), a liquid, or a solid, as is the case with diode lasers, the type you would find in a DVD player.

The laser has a *pumping process*, during which the atomic electrons of the medium are excited to high energy levels. This can be done by means such as electric discharges, flash lamps, or even light from other lasers.

The pumping process increases the energy of the atoms of the medium. Once these atoms are excited, photons injected into the medium cause it to emit other photons: Light shined into the laser medium generates additional light.

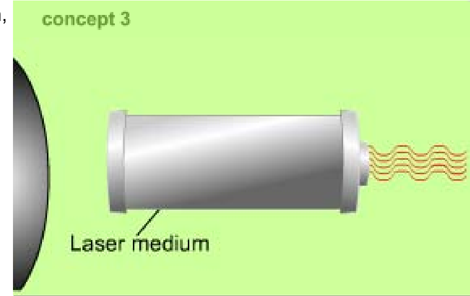
The container has silvered reflective walls, so that photons in the medium are reflected back into it. The process repeats as these photons cause even more in-phase photons to be emitted, and the original light is amplified.

This basic system is illustrated in Concept 3. (Press the refresh button in your browser if you want to see the animation.) The laser medium is inside the container. All the photons reflect off of the mirror on the left, back into the medium, and most of them reflect off the mirror on the right, while some are allowed to pass through. The mirrors form an *optical feedback mechanism*.

This feedback mechanism can be likened to the *audio feedback* that may occur with a sound amplifier. If you have ever winced at the loud squeal as someone experiments with an amplifier, you have heard the unintended consequences of feedback: Too much of the amplified sound is feeding back into the microphone, and is amplified again, and again, until a high-amplification runaway reaction occurs. In a laser, the mirrors reflecting light back into the laser medium provide the feedback. When the system is properly configured, coherent oscillation occurs, and a highly monochromatic, highly directional output beam is created.

Above, we have described the fundamentals of a laser. There remains a basic question, however: Why should the light emanating from the laser be of one color, or frequency? In a laser, you pump energy into the medium, and then use that energy to create an intense beam of light. But why is that light all of the same wavelength, say 633 nm? To explain why, we have to turn to quantum theory.

concept 3

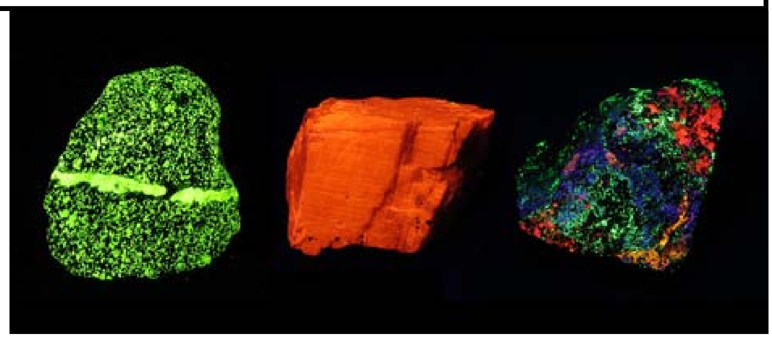


Fundamental laser components
Laser medium: accepts energy, emits light
Pumping process: how electrons get excited
Feedback mechanism: enhances and focuses the signal

41.21 - Laser pumping and stimulated emission

Pumping: Exciting atoms into higher energy levels.

Stimulated emission: An "excited" atom emits a photon when a photon of the right frequency passes by.



Illumination with ultraviolet light makes these minerals fluoresce with characteristic colors. A sample of calcite, in the middle, glows orange.

The core of a laser's functioning is its pumping process, followed by the stimulated emission of radiation. In this section, we describe three types of interaction between radiation and matter. The first, where matter absorbs radiation, is relevant to pumping. The second and third are processes in which matter emits radiation.

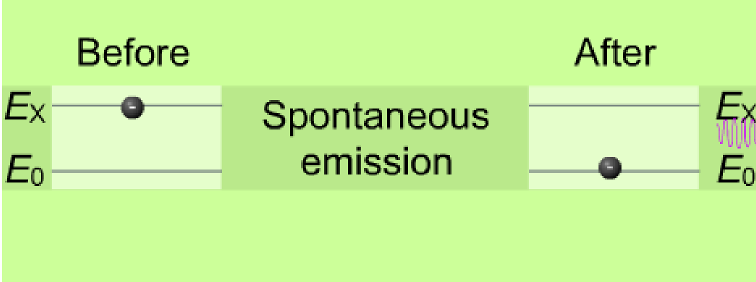
Below, we show absorption. An atom in its *ground state* absorbs a photon and one of its electrons changes states, moving to a higher energy level. These levels are typically described with subscripts; so, for example, the atom might go from energy level E_0 to level E_1.

concept 1

Absorption
Atom absorbs photon, raising energy state

After an atom is excited, how does it return to its ground state? One answer is shown in the illustration below, which illustrates *spontaneous emission*: An atom can spontaneously, with no outside influence, lower its energy state by emitting a photon in a random direction. (This may not seem spontaneous since the atom first had to be "excited," but this is the terminology, and it contrasts with what will be described below.)

concept 2

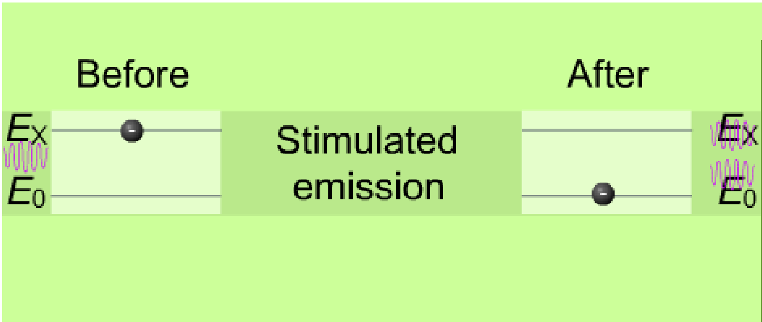


Spontaneous emission
Atom releases photon, lowering energy state

If you shine ultraviolet light upon certain minerals, such as calcite, they *fluoresce*, that is, they glow as long as the energizing ultraviolet radiation continues. A mineral absorbs high-energy photons of radiation and then spontaneously emits photons in a series of steps. These minerals only absorb and emit certain wavelengths of light, due to the quantized nature of their energy levels, so they glow a particular color.

Spontaneous emission is an interesting topic but not the controlled process desired in lasers. If all the laser medium did was to absorb light and then spontaneously re-emit it in random directions, there would be no amplification of the light.

Stimulated emission is the key to lasers. Below, you see an atom that has already been excited to a higher energy state.



concept 3

Stimulated emission
Photon passes near excited atom
Photon triggers emission of 2nd photon
Two photons leave atom

A photon of the appropriate frequency passes close to the excited atom. The result: The atom emits its own photon of the same frequency and returns to its initial energy state. Here is the "gain" produced by lasers: One photon of light causes a net result of two: the process of **l**ight **a**mplification by the **s**timulated **e**mission of **r**adiation. The presence of one photon causes another to be emitted.

We use words like "passes close" and "presence" because photons do not "collide" with atoms during stimulated emission, but they do cause the atom to emit a second photon when the energy of the stimulating photon corresponds to a difference in energy levels allowed in the atom. In sum, with stimulated emission, you start with one photon, and end with two.

To explain further what is occurring in a laser medium, we will use two analogies. The simpler one is mechanistic. Imagine a pool ball that has been raised off the ground and placed on the surface of a flat kitchen table. This increases its gravitational potential energy and corresponds to an atom that has been excited. A second pool ball is rolled at the first and both of them fly off the table. The combined kinetic energy of the two balls as they reach the ground is greater than the original kinetic energy of the ball that was rolled. The collision has "unleashed" the potential energy of the pool ball that was resting on the table.

There is a limitation here, perhaps an issue that concerned you as you considered this metaphor and applied it to a large number of pool balls, in order to make it more similar to the many atoms in the laser medium. The process would not produce a coherent stream of pool balls. If you did this experiment with many pool balls, they would fly off the table in many directions.

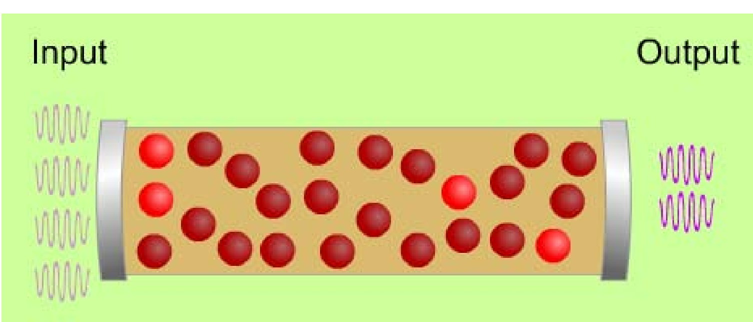
It is better to think of the light as a wave and to consider the phenomenon of resonance, as exhibited by mechanical or electromagnetic waves. Quantum physics states that atoms act like electromagnetic oscillators with particular resonant frequencies. The passing light wave causes the atoms in the medium to begin vibrating in a resonant, coherent relationship. The oscillations are driven by, and in phase with, the stimulating light wave. The atoms respond to the incoming light and reradiate like tiny antennae. The reradiated waves reinforce the waves that cause them.

41.22 - Population inversion

Population inversion: The number of excited atoms is greater than the number of lower-energy atoms.

Stimulated emission is required for lasers to function. For there to be stimulated emission, photons must be passing by and interacting with excited atoms in a medium, from which they can cause coherent photons to be emitted, rather than just passing by atoms whose electrons are at lower energy levels. Not only do the low-energy atoms fail to participate in stimulated emission, but they may also sabotage the process by absorbing photons.

Under normal circumstances, more atoms in a medium will be in lower energy states than in higher energy states. A population inversion is required: The majority of the atoms must be pumped to a higher energy level. When photons strike a medium with an inverted population, stimulated emission is more likely to occur.

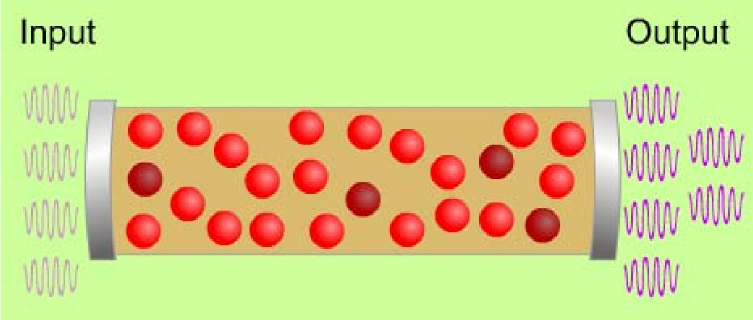


concept 1

An "unexcited" population
Lower-energy atoms outnumber higher-energy atoms
When light shines into medium
· More photons absorbed than emitted
· Light is not amplified

To illustrate the need for an inverted population, we start by showing you the opposite in the diagram above: photons striking a material with a normal population. Some of the atoms are at higher energy levels, but most are not. This is typical of a material that is in thermal equilibrium with its surroundings, such as one at room temperature.

When photons strike the material, most of them are absorbed. This raises the energy levels of some of its atoms, which then spontaneously emit photons in random directions. Occasionally, there is stimulated emission when a photon encounters an atom that already contains an excited electron, but this is rare. This is no way to run a laser!

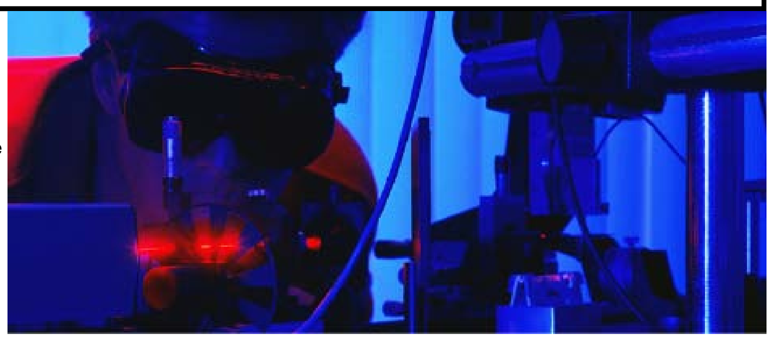


concept 2

Population inversion
Higher-energy atoms outnumber lower-energy atoms
When light shines into medium
· More photons emitted than “fired in”
· Light is amplified

In contrast, the medium above has a population inversion. A photon is far more likely to strike an excited atom, and stimulated emission becomes commonplace. Placing the medium in a reflecting container can further aid this process. Photons will reflect off the container walls and strike other excited atoms, causing a “chain reaction” of stimulated emissions. Energy can be continually added by pumping the medium, ensuring a ready supply of excited atoms.

41.23 - Physics at work: operating a laser



He-Ne laser in a laboratory.

Today, there is a variety of ways to create a population inversion that enables light amplification. To discuss one, let's consider a type of laser that was developed early on, in 1960 and 1961, and remains common, the helium-neon (He-Ne) laser. A glass tube is filled with helium and neon gases. An electric current passes through the tube, and its electrons collide with helium atoms, raising their energy levels from their ground state E_0 to an elevated state E_A.

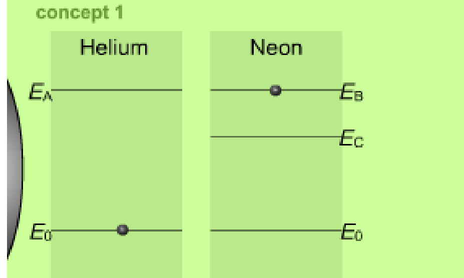
In this state, the helium atoms are said to be at a *metastable* level. They will stay there for a relatively long period of time, as opposed to quickly and spontaneously emitting photons and falling back to a lower energy level. (Relatively long, in this context, means on the order of a thousandth or a ten-thousandth of a second.)

The helium atoms, including those at the metastable level E_A, are continually colliding with neon atoms in the gas mixture. The energy level of a metastable helium atom is very close to that of a high-energy neon atom. In a collision with a ground-state neon atom, a metastable helium atom causes the neon to rise to energy level E_B while the helium atom itself returns to the ground state. This is shown in Concept 1. Neon atoms at E_B are also metastable; they will persist for a relatively long time at this elevated energy level.

Why not just let the electric discharge excite the neon atoms, and skip the helium? Helium is required since neon atoms do not respond readily to electron bombardment by the current that excites the helium atoms.

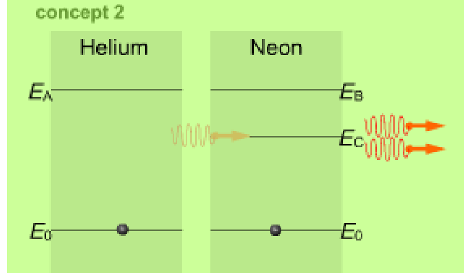
So far, we remain in the dark, so to speak. Many neon atoms have had their energy raised to E_B. The population has been inverted: There are more neon atoms at E_B than at E_0. The next goal is to stimulate the neon atoms to emit photons, dropping to the

concept 1



Laser process: part 1
Energy of helium atoms raised
Energy transferred to neon atoms
Population inversion results

concept 2



Laser process: part 2
Photons enter laser medium
Stimulated emission occurs

· light is amplified
Neon returns to initial state

intermediate energy level E_C in the process. To accomplish this, light is beamed into the laser medium consisting of photons having an energy corresponding precisely to the difference between E_B and E_C. These photons interact with neon atoms whose electrons are at energy level E_B, causing stimulated emissions: The E_B neon atoms each emit a photon of the same frequency, and drop to energy level E_C. The atoms undergo what is called a *laser transition*.

It is important to note that the energy of each stimulating photon must equal the energy difference between E_B and E_C in order for the laser to function. Here is where quantum physics comes into play. A precise understanding of the energy levels of helium and neon atoms is required to design a successful He-Ne laser, as is knowledge of the relationship between the energy and frequency of a photon.

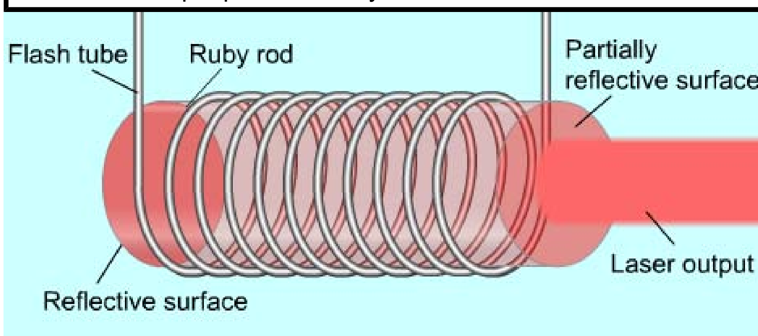
The additional photons that are emitted interact with other energized neon atoms, causing further stimulated emissions. Mirrors on both ends of the laser reflect these photons back into the laser medium, along the axis of propagation. The distance between the mirrors is crucial since it ensures the constructive interference of the light. They require extremely precise fabrication, and reflect more than 99.5% of the photons that hit them.

One last detail: How about the neon atoms that are now at energy level E_C? They rapidly and spontaneously decay to their ground state, emitting photons in all directions in the process. Those photons are in essence “noise” while the stimulated emission photons are “the signal.” The spontaneous emission occurs after about 10^{-8} seconds, ten thousand or so times faster than it occurs with metastable states.

This rapid transition is important. It means that not many neon atoms linger in the E_C state. They quickly return to the E_0 state where collisions with helium atoms can excite them back to E_B, and the lasing process can begin anew. And there is another reason why it is important that few atoms have energy E_C: If there were too many of these around, then they would absorb photons and return to energy E_B, which would thwart the lasing process. (Neon atoms at E_0 cannot absorb photons and rise to either E_B or E_C, since the required energy differences do not match the energy of the photons being used.)

He-Ne lasers are inexpensive and common. The laser tubes can be purchased for less than \$100. These may yield a power output of 1.0 mW when connected to a DC power input of 10 W. This means they are about 0.01% efficient. Such lasers have been used for purposes ranging from supermarket scanners to laser printers, but semiconductor lasers, which are cheaper to fabricate, are rapidly superceding them.

41.24 - Sample problem: ruby laser



The light emitted by a ruby laser has wavelength 694.3 nm, and ruby emits this light when electrons drop from energy level E_2 to E_1 in the ruby's chromium atoms.

What is the energy level difference between E_2 and E_1 in a chromium atom?

Theodore Maimann built the first operational laser in 1960, a flash-pumped ruby laser. Maimann placed a cylindrical ruby rod inside a flash pump. The flash pump excited the chromium atoms in the ruby to a metastable state using an intense burst of light. Then the ions fluoresced back to the ground state, releasing light of wavelength 694.3 nm. Carefully aligned mirrors were placed at either end of the ruby to produce a coherent laser oscillation.

Variables

energy	E
wavelength	$\lambda = 694.3$ nm

What is the strategy?

1. Calculate the energy of a photon of the light emitted from a ruby laser.
2. Use the energy of the light to calculate the energy level difference between E_2 and E_1.

Physics principles and equations

Energy of a photon

$$E = hf$$

Energy level difference

$$\Delta E = E_f - E_i$$

Step-by-step solution

Step	Reason
1. $E = hf$	energy of photon
2. $f = \frac{c}{\lambda}$	frequency, wave speed and wavelength
3. $E = \frac{hc}{\lambda}$	substitute equation 2 into equation 1
4. $\Delta E = E_1 - E_2$ $\lvert\Delta E\rvert = E_2 - E_1$	energy difference
5. $E_2 - E_1 = \frac{hc}{\lambda}$	substitute equation 4 into equation 3
6. $E_2 - E_1 = \frac{(6.63\times10^{-34}\ \mathrm{J\cdot s})(3.00\times10^{8}\ \frac{\mathrm{m}}{\mathrm{s}})}{(694.3\times10^{-9}\ \mathrm{m})}$ $E_2 - E_1 = 2.86\times10^{-19}\ \mathrm{J} = 1.79\ \mathrm{eV}$	evaluate

41.25 - Interactive checkpoint: a ND:YAG laser

In a Nd:YAG laser, the wavelength of the light emitted is 1.064 μm, which is in the infrared region. What is the energy level difference between E_2 and E_1 in this laser?

Answer:

$E_2 - E_1 =$ ______ eV

41.26 - Gotchas

An atom's electrons can exist at any energy level. No, this statement contradicts one of the key tenets of quantum physics. The electrons can only exist at specific energy states, and will not be found with energies between those levels.

A friend says: an electron falls from an energy level of −4.5 eV to −7.2 eV. It emits a photon with 2.7 eV of energy. Has he learned his quantum physics? Yes, he has. The energy of the photon equals the amount of energy given up by the atom.

All photons have the same energy. No, all photons of electromagnetic radiation (for example, light) of a **particular frequency** have the same amount of energy. The energy of a photon increases with the frequency of the radiation.

41.27 - Summary

Something is quantized if it has a smallest, indivisible unit. The opposite of quantized is continuous. The eggs you buy in a carton at the grocery store are quantized; the amount of milk you pour into a glass is effectively continuous.

The Balmer series formula predicts the wavelengths of visible light that are present in the emission spectrum of excited hydrogen gas. When discovered, it revealed a mathematical pattern in the spacing of the spectral emission (or absorption) lines of hydrogen that strongly hinted at an underlying order.

The physicists Max Planck and Albert Einstein showed that radiation is quantized. The quantum of light is called a photon. The energy of a photon of electromagnetic radiation equals Planck's constant times the frequency of the radiation. Einstein cleared up the mystery behind the photoelectric effect by making the assumption that light is quantized.

Niels Bohr developed the basis for the modern-day quantum view of the atom. He stated that the orbits of electrons around the nucleus of an atom are quantized: Electrons can only exist at orbits of specific radii and energy levels. When its electrons jump between levels, an atom emits or absorbs photons whose energy corresponds to the change in electron energy.

Quantum theory is used in the design of semiconductors. Semiconductors are doped − mixed with impurities − to alter the nature of the resistance they offer to currents. Doping creates more mobile electrons and holes than exist in pure semiconductor material.

Key components of transistors include p- and n-type semiconductors. A p-type material has mobile holes that act as charge carriers; an n-type material has mobile electrons. An external potential difference can increase the supply of these charge carrriers near a p-n junction in a transistor, allowing current to flow more readily. However, if the potential difference is changed in magnitude or reversed, it can cause mobile charge carriers to move away from the junction, reducing the flow of current or stopping it altogether.

Laser designers also rely on insights from quantum theory. Some agent such as an electrical discharge excites the laser medium, for example a helium-neon gas mixture. This elevates the medium's electrons to higher energy states. When photons of light are injected into the medium, the excited atoms are stimulated to emit identical, in-phase photons, increasing the flux of photons in the medium and amplifying the light.

For a laser to work, there must be a population inversion: there must be more excited atoms than unexcited ones in the laser medium. If there is no population inversion, then any photons injected into the medium will be absorbed, increasing the energy of the atoms, rather than causing additional photons to be emitted.

Equations

Balmer series

$$\frac{1}{\lambda} = R_{\mathrm{H}}\left(\frac{1}{2^2} - \frac{1}{n^2}\right), \text{ for } n = 3, 4, 5, \ldots$$

Energy of electromagnetic radiation

$$E = hf \qquad \text{(one photon)}$$

$$E = nhf \qquad \text{(multiple photons)}$$

Radii of orbits in a hydrogen atom

$$r_{\mathrm{n}} = n^2 a_0,\ n = 1, 2, 3, \ldots$$

Energy levels in a hydrogen atom

$$E = \frac{-13.6}{n^2} \text{ eV},\ n = 1, 2, 3, \ldots$$

Boltzmann principle

$$\frac{N_2}{N_1} = e^{-(E_2 - E_1)/kT}$$

Chapter 41 Problems

Chapter Assumptions

When converting between electron-volts and joules, assume that

1 eV = 1.60×10^{-19} J

Conceptual Problems

C.1 One estimate for the number of grains of sand on all the beaches on Earth is 5×10^{21} or five thousand billion billion grains. Making the simplifying assumption that sand grains are uniform, give examples where the weight of an amount of sand can be measured without considering the quantized nature of the sand, and cases where this cannot be ignored.

C.2 Suppose two different hydrogen atoms, labeled A and B, have their electrons in state $n = 4$. Atom A emits light when its electron transitions directly from the $n = 4$ state to the $n = 2$ state. Atom B emits light twice, as its electron first transitions from the $n = 4$ state to the $n = 3$ state, then transitions from the $n = 3$ state to the $n = 2$ state. (a) Are the three wavelengths of light the same or different? (b) How does the energy of the photon from atom A compare to the total energy of the two photons from atom B?

(a) i. The wavelengths are the same
ii. The wavelengths are different
iii. Not enough information to tell

(b) i. The energies are equal
ii. The energies are different
iii. Not enough information to tell

C.3 Suppose that for a particular metal, the photoelectric effect occurs with light of wavelength λ_0. With this light incident on a small region of metal, measurements show that electrons leaving the surface have kinetic energies as large as E_0. If the intensity of light is increased without changing the wavelength, (a) how does this affect the number of electrons leaving the surface? (b) What is the effect on the maximum kinetic energy of the ejected electrons?

(a) ○ There are fewer electrons leaving the surface.
○ The number of electrons leaving the surface is unchanged.
○ There are more electrons leaving the surface.

(b) ○ The electrons' maximum kinetic energy is less than E_0.
○ The electrons' maximum kinetic energy is still equal to E_0.
○ The electrons' maximum kinetic energy is greater than E_0.

C.4 Do semiconductors become better or worse electrical conductors at higher temperatures? Explain your answer.

i. Better
ii. Worse
iii. No change
iv. Cannot be determined

C.5 In a *p*-type semiconductor, a current can flow when an electric field is applied and the following happens:

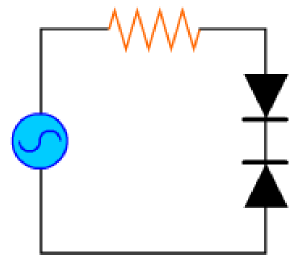
i. There is movement of holes in the valence band.
ii. There is movement of electrons in the valence band.
iii. There is movement of holes in the conduction band.
iv. There is movement of electrons in the conduction band.

C.6 In a *p*-*n* junction, can a hole cross from the *p* side to the *n* side without an electron crossing the junction from the *n* side to the *p* side? Explain your answer.

○ Yes ○ No

C.7 In your own words, explain why current does not readily flow in a circuit loop with a diode that is back-biased.

C.8 Two diodes and a resistor are placed in a single loop circuit connected to an AC generator. The diodes are oppositely oriented, so that when one diode is forward-biased, the other is back-biased. Which phrase best describes the current in the circuit?



i. Nonzero and changing direction with time
ii. Nonzero and constant
iii. Nonzero and changing magnitude with time
iv. Effectively zero at all times

C.9 Two doped semiconductors, one *n*-type and the other *p*-type, are brought together to form a *p*-*n* junction. Before being placed in contact, they are each electrically neutral. They are brought into contact, and reach equilibrium (that is, the diffusion current stops). (a) Is the net charge on the *n*-type semiconductor positive or negative? (b) Is the net charge on the *p*-type semiconductor positive or negative?

(a) ○ Positive ○ Negative
(a) ○ Positive ○ Negative

C.10 Light is amplified in a laser. Consider photons with characteristics that are specific to the laser medium, that is, the laser will emit light at the frequency of the photons. What property of these photons initially increases during the amplification process?

i. Wavelength
ii. Number of photons
iii. Energy of each photon
iv. Frequency

C.11 Does energy need to be continually supplied to maintain a population inversion? Explain your answer.

○ Yes ○ No

Section Problems

Section 2 - Balmer series

2.1 What wavelength light is emitted from a hydrogen atom when the electron transitions from a state where $n = 4$ to one where $n = 2$?

__________ nm

2.2 If a visible photon with wavelength 434 nm is emitted from an atom in a hydrogen gas lamp, what energy state was the atom in before the photon was emitted? That is, what was the number n for that state?

____

Section 4 - Photons

4.1 Visible light consists of photons having energies that range from 2.8×10^{-19} J to 5.0×10^{-19} J. What is the range of photon frequencies in visible light?

__________ Hz up to __________ Hz

4.2 Given the information in the previous problem, what is the range of photon wavelengths in visible light?

lowest wavelength: __________ m
highest wavelength: __________ m

4.3 Gamma ray bursts come from tremendous explosions in faraway galaxies. These bursts are thought to be produced during the end stages of the lives of massive stars, and give off more energy than any other observed astrophysical phenomenon. A gamma photon in such a burst could have a frequency of 10^{21} Hz. A typical visible photon has a wavelength of 550 nm. Roughly how many times more energetic is this gamma photon than a typical visible photon?

i. 10
ii. 1000
iii. 1000000
iv. 1000000000

Section 6 - Photoelectric effect

6.1 Electrons can be ejected from a clean silver surface using light with wavelengths as large as 262 nm. What is the work function for silver?

__________ eV

6.2 Potassium has a work function of 2.21 eV. What is the maximum kinetic energy of electrons escaping from a sample of potassium if photons of frequency 6.00×10^{14} Hz strike it?

__________ J

6.3 In a demonstration of the photoelectric effect, electrons leave a calcium surface with velocities as great as 2.0% of the speed of light. If the work function for calcium is 3.2 eV, what is the wavelength of the incident radiation? Relativistic effects may be neglected.

__________ nm

6.4 Platinum has a work function of 6.4 eV. Radiation with wavelength 150 nm and intensity 1.3 W/m^2 is directed at a platinum surface. (a) Find the maximum kinetic energy of the ejected electrons. (b) If 1.0% of the photons striking a 1.0 cm^2 area of platinum eject electrons, how many electrons are emitted per second from that part of the surface?

(a) __________ J
(b) __________ electrons per second

Section 9 - Bohr atom

9.1 If a 12.1 eV photon was absorbed by a hydrogen atom initially in its ground state, which then emits 2 photons, what are the emitted photon wavelengths?

shorter wavelength: __________ nm
longer wavelength: __________ nm

Section 13 - Conduction in solids

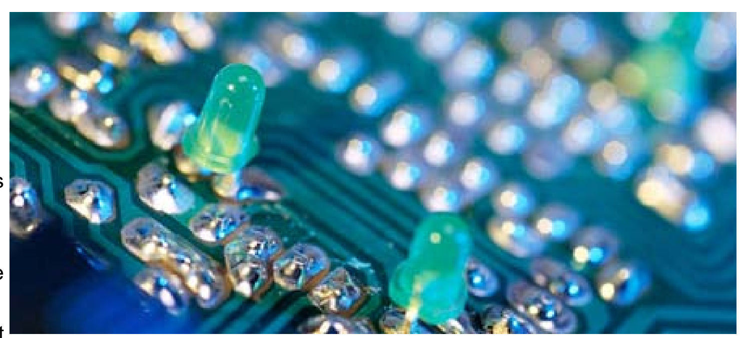
13.1 What is the longest wavelength that a photon can have to promote an electron from the top of the valence band to the bottom of the conduction band in silicon? Assume a band gap of 1.12eV for silicon.

__________ m

13.2 What is the longest wavelength that a photon can have to promote an electron from the top of the valence band to the bottom of the conduction band in silicon dioxide? Assume an energy gap of 8.0 eV between the valence and conduction bands.

__________ m

Section 16 - p-n junction

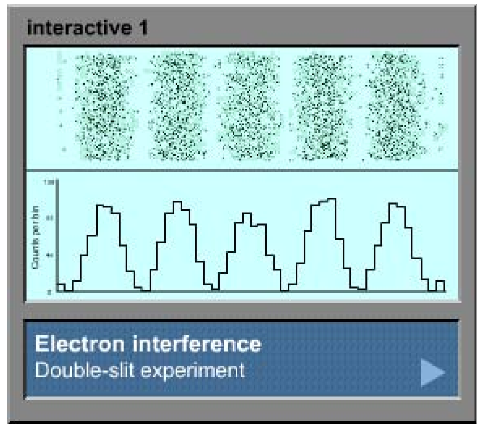


16.1 Diodes in forward-biased circuits emit electromagnetic radiation. Specifically, when an electron from the conduction band in the *n*-type semiconductor of the diode fills a hole in the valence band of the *p*-type semiconductor of the diode, it emits a photon whose energy is equal to the difference in energy between the two bands. The energy required for this emission is provided by whatever is driving the electric current in the circuit.
The accompanying photograph shows a circuit containing green *light emitting diodes* (LEDs). LEDs actually look a lot like computer chips, but they are enclosed in protective plastic lenses like the ones you see here so that they will project a bright light. Light emitting diodes are made from non-silicon semiconductors such as gallium, and the energy difference between their conduction and valence bands is such that the photons they emit have wavelengths in the visible light spectrum. LEDs are used for applications ranging from jumbo TV screens in sports stadiums to the digits of glow-in-the-dark clocks.
The LEDs in the photograph emit green light with a wavelength of 520 nm. What is the energy difference between their conduction and valence bands?

__________ eV

chapter 42 Quantum Physics Part Two

42.0 - Introduction



Richard Feynman (1918 − 1988) was an American physicist best known to the general public for his leading role in the commission that investigated the destruction of the space shuttle Challenger in 1986. His bestselling memoirs delighted millions with his irreverence and tales of bongo drum-playing.

In the physics community, Feynman was known for his insight into quantum mechanics and the brilliance of his lectures. The "Feynman Lectures in Physics," still in print today, show his timeless ability to clearly and memorably explain fundamental aspects of physics.

We start this chapter by giving you a chance to conduct a thought experiment that Feynman described as "impossible, absolutely impossible, to explain in any classical way, and which has in it the heart of quantum mechanics." In this thought experiment, Feynman emphasized the shortcomings of considering matter solely as composed of particles, and the need to adopt a vision where matter has both a particle and a wave nature. This is known as the wave-particle duality.

Feynman first asked the reader to consider classical particles, that is, those for which the physics of Newton completely predicts the motion. Feynman used bullets in his experiment; we will be slightly more pacific and ask you to imagine throwing baseballs at a picket fence. The picket fence is missing two of its slats. The missing slats provide gaps that the baseballs can pass through. The balls will stop when they strike a wall behind the fence.

If you conducted this experiment yourself, you would find that most of the baseballs hit the wall directly behind the missing slats. A few baseballs would hit slightly to the sides of those areas, corresponding to the balls that went through the gaps at an angle. You would see two piles of baseballs that accumulated behind the gaps in the fence.

Feynman used this first thought experiment to remind his audience how they expect the world to "work" based on everyday experience. This experiment provided the contrast with another experiment, one you can conduct for yourself using the simulation to the right.

In the simulation, you are conducting a similar experiment, but with particles of far smaller scale. Instead of baseballs being fired, electrons are fired one at a time toward a barrier with two slits. The slits are very narrow and very close together.

The electrons pass through the slits and reach a screen, which is a photographic material that records where they land. The electrons cannot be observed as they move from the source to the barrier, or from the barrier to the screen. Only the final position is marked, by using black dots.

Press FIRE to launch a single electron. Then, hold the FIRE button down to fire a stream of electrons so you can see the pattern of where they land.

Look at the pattern of where the electrons accumulate. Instead of accumulating in two piles, like the baseballs did in Feynman's first thought experiment, you see regions where many electrons accumulate, alternating with regions where very few electrons land. The pattern should remind you of the dark and light fringes that are created on a screen when light shines through a pair of slits.

In the theory of optics, physicists explain the pattern of light and dark fringes by modeling light as a wave, with wave-like properties such as frequency and wavelength. This simulation shows that something which you have solely considered to be a particle, an electron, also displays wave-like properties in a similar experiment. This is the essential point of Feynman's second thought experiment: Particles such as electrons have wave-like properties. A single particle can travel from a source to a screen and demonstrate interference effects due to the presence of the two slits. The wave that is associated with a moving particle is called a *matter wave*.

Now you can use the simulation again to see another fundamental aspect of quantum mechanics. Reset the simulation. Press the FIRE button a few times and note the locations of the first three or four electrons. Then press RESET again, fire a few more electrons, and note their locations. In your two experiments, did the first three or four electrons show up at the same location each time?

The answer to the question is "no". In both this simulation, and in the real experiments that this simulation is recreating, the location of a single electron cannot be predicted. Although the overall pattern of light and dark fringes as you fire hundreds of electrons can be predicted, where a particular electron will strike the screen cannot be determined in advance.

This is the second point of the animation: The locations of the electrons can be stated in terms of likelihoods, like the probability of drawing an ace from a deck of cards. The pattern of light and dark fringes provides a "map" of where an electron is likely to strike the photographic paper behind the slits. However, you cannot predict in advance where any one electron will land, any more than you can state with certainty that one particular card will be an ace.

You are witnessing a major point of quantum mechanics in this simulation. Scientists have performed the experiment with electrons, and the results have been exactly as depicted. Electron after electron can be fired through slits, and the interference of their matter waves will create a pattern like the one you see here.

The simulation shows the wave-like properties of particles. Scientists like Einstein had also postulated the particle-like nature of light, which had been considered a wave. This chapter will start with that topic, the particle nature of electromagnetic radiation, before returning to the topic that the experiment in this section illustrates.

Compton effect: Considering x-rays as composed of particles explains why their frequency diminishes as they are scattered by a material.

Einstein stated that light was composed of indivisible packets of energy. His theory was rooted in experiment. By treating light as a particle he could explain the photoelectric effect and make successful predictions: Calculating the energy of the photons that make up light of a particular color allowed Einstein to predict whether shining that color light on a given material would cause it to emit electrons.

This opened up another question: Did light have other particle properties, such as momentum? In 1916, Einstein derived an equation quantifying a photon's momentum. As shown in Equation 1, the momentum equals Planck's constant divided by the photon's wavelength. However, it remained to be shown experimentally that photons actually had momentum. Some scientists believed that the photoelectric effect might be explained by updating the wave theory of light, but if light could be shown to have momentum then it would be hard to deny its particle-like nature.

Physicists such as W.H. and W.L. Bragg − a famous British father-and-son pair − had been using x-rays to analyze the structure of matter, particularly crystals. The small wavelengths of x-rays, compared to the atomic spacing in the crystals, meant that the scattered radiation would exhibit diffraction patterns, that is, the radiation would be strong in particular directions. The diffraction pattern could be used to analyze the regular, periodic layout of atoms within crystals, and this technique was later used to deduce the double-helix structure of DNA. Since diffraction is a property only of waves, you might think that x-ray diffraction experiments could not provide any support for a particle view of electromagnetic radiation − but as it turned out, they did.

Scientists experimenting with x-rays scattering from target materials started to observe some disturbing data. To explain their discomfort, we first have to explain what they expected.

Their expectations were based on the view of light as an electromagnetic wave, consisting of oscillating electric and magnetic fields. In this classical picture, when an electromagnetic wave encounters an atom, it causes the atom's electrons to oscillate at the same frequency as the wave. Accelerating charges emit radiation, so these oscillating charges then radiate their own electromagnetic waves. This process is called *scattering*; some waves would be re-emitted in the same direction as the initial wave, while others would be re-emitted in other directions.

The critical point to focus on is the frequency of the re-radiated waves. Classical electromagnetic theory, used by scientists like the Braggs, predicted that the outgoing waves should have the same frequency as the incoming waves. Scattering would change the direction of the incoming radiation, but scientists were confident that it should **not** change its frequency.

Alas, observations did not accord with theory. Scientists observed that atoms being subjected to x-rays were re-emitting radiation of lower frequency than the initial radiation, and the effect depended on the angle at which the radiation was scattered. Something was diminishing the frequency (and the energy) of the scattered radiation.

At first, the data were not taken seriously. However, in 1923 the American physicist A. H. Compton published two papers that argued conclusively that if radiation were quantized, one should expect the frequency to be reduced. The problem was not with the experimental data; the problem was with a theory that regarded light solely as a wave.

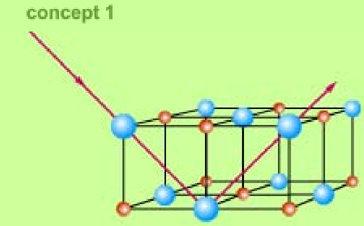
How did Compton explain the discrepancy? He assumed that the electromagnetic radiation was made up of photons that had momentum. He proposed that instead of picturing a light wave shaking electrons up and down, scientists ought to picture the interaction as akin to a collision between two particles, a photon and an electron. Like Einstein, he was asking his peers to expand their conception of radiation to include properties usually associated with particles, such as momentum.

He then used classical mechanics, applying the laws of conservation of energy and momentum to the collision. He applied the same principles that would be applied to a collision of two billiard balls. (His analysis had to be more complex than that for two balls, because he had to relate the quantized photon energy to wavelength and frequency, and relate its energy and momentum.)

Compton stated that when the photon in question collides with an electron belonging to an atom in the target, the electron gains some kinetic energy from the collision. Energy must be conserved, and the photon loses that same amount of energy. The energy of a photon equals hf, the product of Planck's constant and its frequency. When its energy is diminished in the scattering process, so too is its frequency.

Not only did Compton's work explain the reduction in frequency, but his analysis of the collision also correctly relates the direction of the

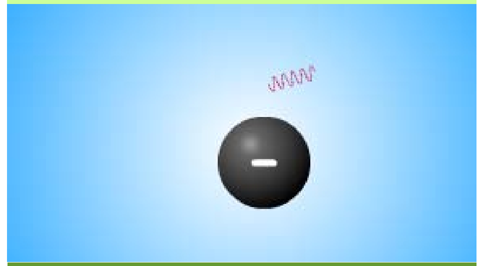
concept 1



Electromagnetic radiation as wave

When an x-ray met a crystal:
- Scientists anticipated its frequency would stay the same
- They were wrong

concept 2



Compton effect

Consider x-ray as a photon (particle)
A photon "collides" with an electron in the crystal
Photon loses energy in collision
Less energy means lower frequency ($E = hf$)

equation 1

Momentum of a photon

$$p = \frac{h}{\lambda}$$

p = momentum
λ = wavelength
h = Planck's constant, 6.63×10^{-34} J·s

scattered x-rays to their change in frequency. When a moving particle strikes a stationary target, its change in momentum and the angle at which it scatters are related. Consider the photon-electron collision. If the direction at which the photon comes out is virtually unchanged from its original direction, then its change in momentum is small, and very little momentum (and energy) will be given to the electron. In other words, the two particles just suffered a glancing blow.

When Compton observed x-ray photons whose direction was barely changed, he saw that their frequency was also practically unchanged. Again, the basics of collisions held true.

As the angle between the emitted and incident radiation increased, the change in momentum increased. More momentum (and energy) is transferred to the electron. Radiation that was *backward scattered* suffered the largest reduction in frequency.

Compton's work with x-rays confirmed that electromagnetic radiation possesses momentum, a property that had been classically associated with particles. It was increasingly hard to argue that light should only be considered a wave when it could be demonstrated that it had momentum, and its interaction with matter could be modeled using a classical explanation of collisions. It became necessary to admit that under certain conditions the wave nature of light is observed while in different experiments the particle nature is needed to explain the results.

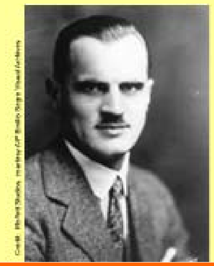
42.2 - Derivation: the Compton effect equation

At the right, we show the equation for the Compton effect. The change in wavelength of the radiation is stated as a function of Planck's constant, the mass of the electron, and the angle between the photon's initial path and its path after the collision.

The factor $h/m_e c$ in the equation is known as the *Compton wavelength* of the electron, and has dimensions of length.

We will derive the equation, along the way showing an important relationship between the total energy, momentum and rest energy of a particle with mass. The derivation also enables us to discuss the source of the equation for a photon's momentum.

equation 1



Compton's equation

$$\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$$

λ = wavelength
h = Planck's constant
m_e = mass of electron
θ = angle between initial and final paths of photon
$\frac{h}{m_e c}$ = Compton wavelength of electron

Part 1

Before considering the collision between photon and electron, we first derive a very useful equation that relates the total energy, momentum, and rest energy of a particle.

Variables

total energy of particle	E
mass of particle	m
speed of particle	u
momentum of particle	p
speed of light	c

Strategy

1. Start with the relativistic expressions for the momentum and total energy of a particle.
2. Divide the two equations to obtain the particle speed in terms of total energy and momentum.
3. Substitute the expression for u into the equation for relativistic momentum, then solve for the total energy.

Physics principles and equations

The equation for the relativistic momentum of a particle

$$p = \gamma m u = \frac{mu}{\sqrt{1 - \frac{u^2}{c^2}}}$$

The equation relating the total energy of a particle to its rest mass

$$E = \gamma m c^2 = \frac{mc^2}{\sqrt{1 - \frac{u^2}{c^2}}}$$

Planck's equation for the energy of a photon

$$E = hf$$

Speed of an electromagnetic wave

$$c = \lambda f$$

Step-by-step derivation

Step	Reason
1. $\frac{p}{E}=\frac{u}{c^2}$	divide momentum equation by energy equation
2. $u=\frac{p}{E}c^2$	solve for u
3. $p=\frac{mu}{\sqrt{1-\frac{u^2}{c^2}}}$	relativistic momentum
4. $p^2=m^2u^2+\frac{p^2u^2}{c^2}$	square both sides and rearrange terms
5. $p^2=m^2\left(\frac{p}{E}c^2\right)^2+\frac{p^2\left(\frac{p}{E}c^2\right)^2}{c^2}$	substitute speed from step 2 into the equation of step 4
6. $E=\sqrt{(pc)^2+(mc^2)^2}$	solve for the total energy E

The expression above for the total energy of a particle was derived using relativistic expressions for energy and momentum. It can be applied to any particle. We will use the equation to derive the momentum of the photon. The equation above simplifies since a photon has zero mass. With that fact, we can use Planck's equation for energy and the relationship of wave speed, wavelength and frequency to derive the equation for the *momentum of a photon*.

Step	Reason
1. $E=\sqrt{(pc)^2+(mc^2)^2}$	equation derived above
2. $E_{\text{photon}}=p_{\text{photon}}c$	equation applied to photon, which has zero mass
3. $E=hf$	Planck's equation
4. $hf=p_{\text{photon}}c$	substitute Planck's equation into step 2
5. $f=\frac{c}{\lambda}$	frequency stated in terms of wave speed and wavelength
6. $\frac{h}{\lambda}c=p_{\text{photon}}c$	substitute equation for frequency into step 4
7. $p_{\text{photon}}=\frac{h}{\lambda}$	speed of light cancels

This important result relates the magnitude of the photon's momentum to its wavelength. This relationship will be needed to complete the derivation of the Compton effect equation. We are now ready to analyze the collision between photon and electron.

Part 2

Variables

total energy	E
mass of electron	m_e
momentum	p
wavelength of light	λ
angle between initial, final paths of photon	θ

The subscripts e and γ denote the electron and photon, respectively, and i and f indicate the initial (pre-scattering) or final (post-scattering) value.

Strategy

1. Use the principle of conservation of energy and the expression for the total relativistic energy of a particle developed in part 1 to obtain an expression for the square of the final momentum of the electron.

2. Use conservation of momentum to develop another expression for the square of the final momentum of the electron.

3. Equate the two expressions for the square of the final electron momentum to eliminate that variable. Rearrange the result to obtain the Compton effect equation that relates the initial and final photon wavelengths.

Physics principles and equations

In addition to the conservation laws indicated above, we will use the fact that momentum is a vector quantity.

We are assuming that the electron is at rest before the collision.

Mathematics principles

The dot product of two vectors is related to the angle between them, θ.

$$\mathbf{A}\cdot\mathbf{B} = AB\cos\theta$$

The square of the magnitude of a vector can be obtained by taking the dot product of the vector with itself.

$$A^2 = \mathbf{A}\cdot\mathbf{A}$$

Step-by-step derivation

We will first use the conservation of energy to develop an expression for the square of the final electron momentum.

Step	Reason
1. $E_{i\gamma}+E_{ie}=E_{f\gamma}+E_{fe}$	conservation of energy
2. $E_{ie}=m_ec^2$	electron initially at rest
3. $E_{fe}=m_ec^2+E_{i\gamma}-E_{f\gamma}$	combine steps 1 and 2
4. $\sqrt{(p_{fe}c)^2+(m_ec^2)^2}=m_ec^2+p_{i\gamma}c-p_{f\gamma}c$ $c\sqrt{p_{fe}^2+m_e^2c^2}=m_ec^2+p_{i\gamma}c-p_{f\gamma}c$	substitute energy expressions from part 1
5. $p_{fe}^2=(m_ec+p_{i\gamma}-p_{f\gamma})^2-m_e^2c^2$	solve for square of final electron momentum

Next we use the conservation of momentum to develop a second expression for the square of the final electron momentum.

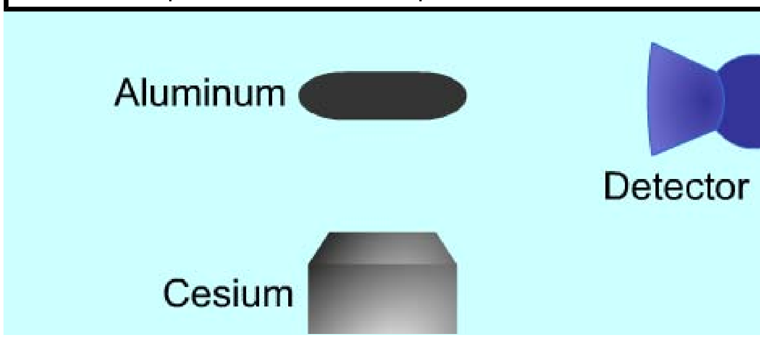
Step	Reason
6. $p_{fe}^2=\mathbf{p}_{fe}\cdot\mathbf{p}_{fe}$	property of vector dot product
7. $\mathbf{p}_{ie}+\mathbf{p}_{i\gamma}=\mathbf{p}_{fe}+\mathbf{p}_{f\gamma}$ $\mathbf{p}_{fe}=\mathbf{p}_{i\gamma}-\mathbf{p}_{f\gamma}$	conservation of momentum
8. $p_{fe}^2=(\mathbf{p}_{i\gamma}-\mathbf{p}_{f\gamma})\cdot(\mathbf{p}_{i\gamma}-\mathbf{p}_{f\gamma})$ $p_{fe}^2=p_{i\gamma}^2+p_{f\gamma}^2-2p_{i\gamma}p_{f\gamma}\cos\theta$	substitute step 7 into step 6; take dot products

Equating the two expressions for the square of the final electron momentum allows us to eliminate the electron momentum from the final equation.

Step	Reason
9. $(m_e c + p_{i\gamma} - p_{f\gamma})^2 - m_e^2 c^2 = p_{i\gamma}^2 + p_{f\gamma}^2 - 2p_{i\gamma}p_{f\gamma}\cos\theta$ $p_{i\gamma}p_{f\gamma}(1-\cos\theta) = m_e c\,(p_{i\gamma} - p_{f\gamma})$	equate expressions in steps 5 and 8
10. $\frac{1}{p_{f\gamma}} - \frac{1}{p_{i\gamma}} = \frac{1}{m_e c}(1-\cos\theta)$	rearrange
11. $\frac{\lambda_f}{h} - \frac{\lambda_i}{h} = \frac{1}{m_e c}(1-\cos\theta)$ $\lambda_f - \lambda_i = \frac{h}{m_e c}(1-\cos\theta)$	use relationship between momentum and wavelength for photon

The success of this equation in correctly stating the relationship between the change in wavelength and the scattering angle gave physicists great confidence that the principles employed in the derivation were correct.

42.3 - Sample Problem: the Compton effect



High energy photons emerge from radioactive cesium with an energy of 0.662 MeV. They are scattered by the loosely bound electrons in an aluminum target and measured by a detector at 90 degrees from the initial direction. What is the expected energy of these photons?

Variables

energy of incoming photon	$E_i = 0.662$ MeV
energy of scattered photon	E_f
wavelength of incoming photon	λ_i
wavelength of scattered photon	λ_f
electron rest energy	$m_e c^2 = 0.511$ MeV
angle between initial and final photon paths	$\theta = 90°$

What is the strategy?

1. Relate the energy of each photon to its wavelength.
2. Use the Compton effect equation to determine the change in wavelength.
3. Solve for the energy of the scattered photon.

Physics principles and equations

The relationship between the energy of a photon and its wavelength is given by the Planck relation.

$$E = hf = \frac{hc}{\lambda}$$

The change in photon wavelength is given by the Compton effect equation.

$$\lambda_f - \lambda_i = \frac{h}{m_e c}(1-\cos\theta)$$

Step-by-step solution

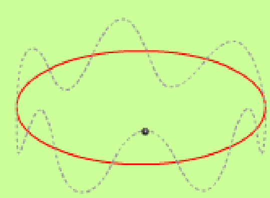
Step	Reason
1. $E=\dfrac{hc}{\lambda}$ $\lambda=\dfrac{hc}{E}$	photon energy
2. $\lambda_f-\lambda_i=hc\left(\dfrac{1}{E_f}-\dfrac{1}{E_i}\right)$	use step 1 to express the change in wavelength
3. $\lambda_f-\lambda_i=\dfrac{h}{m_e c}(1-\cos\theta)$	Compton effect equation
4. $hc\left(\dfrac{1}{E_f}-\dfrac{1}{E_i}\right)=\dfrac{h}{m_e c}(1-\cos\theta)$	equate expressions for change in wavelength
5. $\dfrac{1}{E_f}-\dfrac{1}{E_i}=\dfrac{1}{m_e c^2}(1-\cos\theta)$ $\dfrac{1}{E_f}=\dfrac{1}{m_e c^2}(1-\cos\theta)+\dfrac{1}{E_i}$ $E_f=\left[\dfrac{1}{E_i}+\dfrac{1}{m_e c^2}(1-\cos\theta)\right]^{-1}$	solve for the final energy
6. $E_f=\left[\dfrac{1}{0.662\ \text{MeV}}+\dfrac{1}{0.511\ \text{MeV}}(1-\cos 90°)\right]^{-1}$	substitute values
7. $E_f=0.288\ \text{MeV}$	evaluate

42.4 - Matter waves

concept 1

De Broglie
Matter is particle with wave-like properties

concept 2



De Broglie and the electron

After Compton's observations, physicists were forced to confront the fact that light, which had been thought of as solely an electromagnetic wave, also has properties of a particle. This new viewpoint enabled them to understand the experimental data they were confronting.

However, there still remained many enigmas. For instance, Bohr had constructed his model of the hydrogen atom, which successfully predicted observed emission and absorption spectra. Bohr proposed quantized energy states for the electron, starting with a classical view of the electron as a negatively charged, point-sized particle circling the positive nucleus, and using Planck's work. However, Bohr was not able to demonstrate why his model was correct.

In 1923, a French doctoral student named Louis de Broglie proposed a simple idea to help to rescue the physicists from their intellectual tar pit. Recognizing that light has both wave and particle properties, he reasoned that nature is symmetrical, and that the same is true for matter. De Broglie asserted that particles such as electrons have both particle and wave properties.

To quote de Broglie: "…I had a sudden inspiration. Einstein's wave-particle dualism was an absolutely general phenomenon extending to all physical nature…"

Although it was a simple idea, it propelled the next revolution in physics.

De Broglie conceived of the electron in an atom as a standing "matter wave" vibrating around the nucleus. In other words, the electron is "smeared out" instead of being a single point-sized particle. Let's consider the implications of representing an electron with such a wave, as shown in Concept 2.

As you may recall, a standing wave results from waves that interfere with one another. For there to be constructive interference, peak must meet peak, and trough must meet trough. In contrast, if at a given location, the peak of one wave meets the trough of another, the result is destructive interference − a flat line, in essence.

Here, the string is looped into a circle, and the length of the string is the circumference of the circle. A wave begins its circular path around the string, and when it makes a loop, it meets up with itself.

If as it makes a second journey around, peak meets peak, the result will be a standing wave. On the other hand, if peak meets trough, the wave will cancel. The relationship between the circumference of the circle and the wavelength determines whether there is constructive interference. In Concept 2, we show an example where there is constructive interference.

The wave shown to the right does not represent the actual path of the electron through space. It is a matter wave − a way to visualize the likelihood of finding the electron at a given location. If the amplitude of the wave is zero, then the likelihood is zero, and there will be no electron.

This provides one piece of the puzzle. Only certain wavelengths are possible for a given orbit, but that doesn't prevent any orbit being possible − it just dictates possible wavelengths.

The other piece of the puzzle comes from considering the angular momentum of the electron. De Broglie showed why Bohr's quantization argument could be justified by considering the angular momentum of the electron and the relationship between momentum and wavelength.

orbits
Wave-like properties explain quantized orbits

equation 1

Wavelength of a matter particle

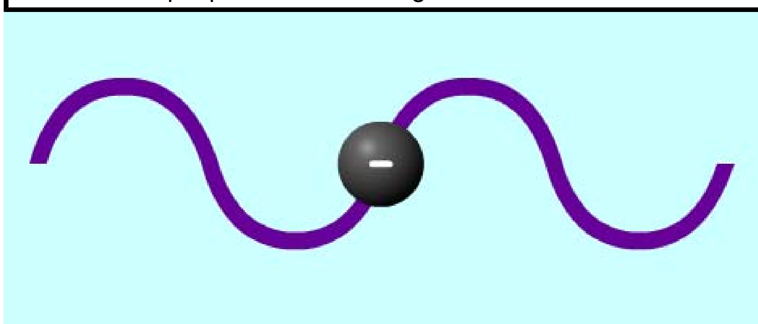
$$\lambda = \frac{h}{p}$$

λ = wavelength
p = momentum
h = Planck's constant, 6.63×10^{-34} J·s

To quantify the wavelength of a matter particle such as an electron, de Broglie proposed that the same equation that describes the momentum of a photon, $p = h/\lambda$, could also be applied to matter. A particle's wavelength, sometimes called the de Broglie wavelength, is related to its momentum by $\lambda = h/p$, as shown in Equation 1.

De Broglie's insight provided a crucial step in the understanding of the atom. He and other physicists used the idea to write down equations for the standing waves corresponding to any particle confined to a small space. The form of the matter waves (also known as the wavefunction of the particle) leads to predictions about the behavior of the particle. This is the principle behind quantum mechanics, also called wave mechanics.

42.5 - Sample problem: wavelength of an electron



What is the wavelength of an electron's matter wave when it has a kinetic energy of 1000 eV?

If there is a wave aspect to matter − "matter waves" − why do we not see this every day? Why do we not observe the wave aspect of, say, a golf ball's motion through the air? The answer is that the wavelength is exceedingly small for everyday objects moving at ordinary speeds.

This problem asks you to quantify the wavelength of a moving electron with a certain kinetic energy. Then we will briefly use this example to discuss the wavelengths of objects of larger mass.

Variables

electron kinetic energy	KE = 1000 eV
electron wavelength	λ
electron momentum	p
electron mass	$m = 9.11\times10^{-31}$ kg

What is the strategy?

1. Determine the momentum of the electron.

2. Use the de Broglie hypothesis to determine the wavelength of the electron.

Physics principles and equations

The (non-relativistic) kinetic energy of a particle is

$$KE = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

The electron wavelength is related to its momentum by

$$\lambda = \frac{h}{p}$$

Step by step solution

Step	Reason
1. $KE = \frac{p^2}{2m}$	kinetic energy
2. $p = \sqrt{2mKE}$	solve for the momentum
3. $\lambda = \frac{h}{p}$	de Broglie hypothesis
4. $\lambda = \frac{h}{\sqrt{2mKE}}$	substitute momentum from step 2
5. $\lambda = \frac{6.63 \times 10^{-34}\,\mathrm{J \cdot s}}{\sqrt{2(9.11 \times 10^{-31}\,\mathrm{kg})(1000\ \mathrm{eV})(1.60 \times 10^{-19}\,\mathrm{J/eV})}}$	substitute values
6. $\lambda = 0.0388\ \mathrm{nm}$	evaluate

Note that this wavelength is thousands of times smaller than the wavelength of visible light.

We can calculate and discuss the wavelength of everyday objects, such as a golf ball, using the equation in step 3. The wavelength is Planck's constant divided by the momentum of the object. Even for objects that are not very massive by everyday standards, the wavelengths are incredibly tiny, and do not have a visible effect on the motion of these objects. For example, a one-microgram dust particle drifting along at one mm/s has a wavelength that is trillions of times less than that of the electron above, which is on the order of hundredths of a **nanometer**. To obtain a wavelength that we could observe, say one millimeter, a one-kilogram object would have to be moving on the order of 10^{-30} m/s. This slow speed makes observing its wavelength a bit difficult!

42.6 - Observing matter waves

When light is shone through a pair of small slits, an interference pattern results. We show a light interference pattern in Concept 1.

If electrons can act like waves, they should also display interference patterns, and they do. The slits used should be of a width comparable to the wavelengths of the electrons in the experiment, which are moving at a speed such that their wavelength is on the order of 10^{-10} meters. You see the interference pattern caused by sending electrons through slits in Concept 2.

At the time that de Broglie proposed his theory of matter waves, it was not possible to make slits small enough to demonstrate electron diffraction. However, in 1927 two physicists named Clinton Davisson and Lester Germer inadvertently produced electron diffraction using a crystal of nickel. The spacing between atoms in the crystal happened to be on the order of the electron wavelength, causing the electrons' matter waves to interfere.

After witnessing this and the diffraction of other particles such as neutrons and whole hydrogen atoms, scientists began to take the wave-like nature of particles for granted, or perhaps better put, to marvel at it as a fact of nature. They called it the *wave-particle duality*.

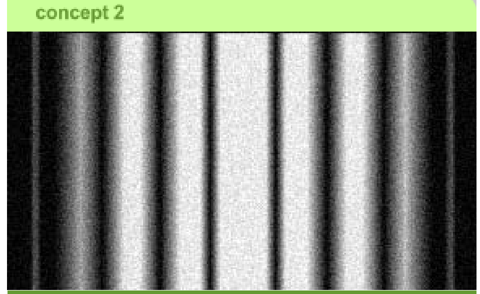
Once they finished marveling, they also concluded that they could take advantage of the wave properties of matter. Wave diffraction imposes a limit on how small of an object can be resolved when it is probed with radiation of a certain wavelength. For example, an ordinary microscope uses visible light and glass lenses, and cannot resolve objects much smaller than 10^{-6} m, which is on the order of the wavelength of visible light.

To gain increased resolution, a *transmission electron microscope* (TEM) employs electrons instead of light. The wavelengths of those electrons are about ten thousand times smaller than that of light, which allows the TEM to resolve objects down to a size of about 10^{-10} m. Just as the light in an ordinary microscope passes through a sample that is fixed on a glass slide, the electron beam (think of it as a wave) passes through the thin sample on the way to a detector. An ordinary microscope uses a glass lens to focus light rays; the TEM uses magnetic fields to focus the charged electron beam.

concept 1

Finding the wave in light
Light's wave-like properties visible in interference patterns

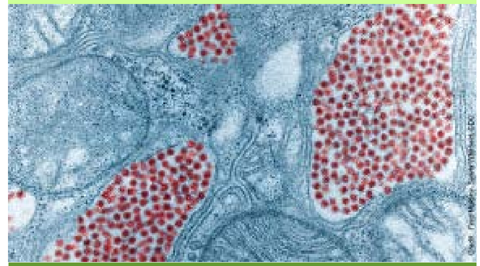
concept 2



Matter waves
Electron interference pattern
Pattern also created by projecting electrons at crystal

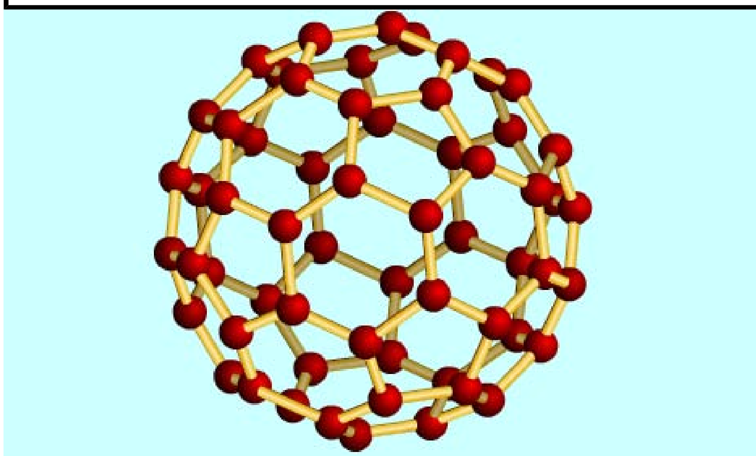
In Concept 3 is an artificially-colored image captured by a transmission electron microscope. It shows a salivary gland of a mosquito infected by the Eastern equine encephalitis virus (red dots). The individual viruses are about 60 nanometers in diameter, and even smaller details than this are visible in the image. In comparison, the best optical microscopes can only resolve details as small as 200 nanometers.

concept 3



Transmission electron microscope
Takes advantage of small wavelengths of electrons
Resolves details at a scale of 10^{-10} meters

42.7 - Interactive checkpoint: buckyball interference



You wish to observe the wave-like properties of a relatively large particle: a molecule of C_{60} (a "buckyball"). You heat a powder of C_{60} in an oven and the molecules exit it with a velocity distribution that peaks at 225 m/s. What is the wavelength of the molecules that are moving at the most probable velocity? If these buckyballs pass through a pair of slits whose center-to-center separation is 50.0 nm, what will be the angle to the first off-center maximum in the interference pattern?

"Buckyball" is the nickname for C_{60}, or buckminsterfullerene. It is a soccer ball shaped molecule named after the architect Buckminster Fuller, who designed the geodesic domes the molecule resembles. Physicists have been able to observe diffraction of C_{60} as well as C_{70}, a less stable type of fullerene. The wave properties of even larger molecules, such as porphyrin (a molecule found in plants) and $C_{60}F_{48}$, have also been observed.

C_{60} is composed of 60 carbon molecules, each with a mass of 12 amu. This means its overall mass is approximately
$60 \times 12 \times (1.67\times10^{-27}\text{ kg}) = 1.20\times10^{-24}\text{ kg}$.

Answer:

$\lambda =$ ______ m

$\theta =$ ______ °

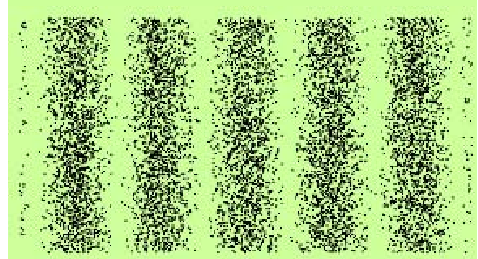
42.8 - Matter waves are probabilistic

De Broglie correctly asserted that electrons and other particles can be described as matter waves. Physicists, being quantitative types, wanted to know: How could they mathematically describe the location of a wave-particle like an electron?

In classical mechanics, when the net force exerted on a particle of a certain mass is known, the equation $F = ma$ is used to calculate the acceleration of the particle. If the initial position and velocity are also known, then other equations can determine the particle's location and velocity at any time.

Quantum physicists found that applying classical equations to atomic-sized particles led to paradoxes. For instance, the equations could be interpreted to predict that a quarter photon, or half an electron, should be present at a given location. This was in contradiction not only to the tenets of quantum physics, but to experiments, which showed that photon and electrons were indivisible.

concept 1

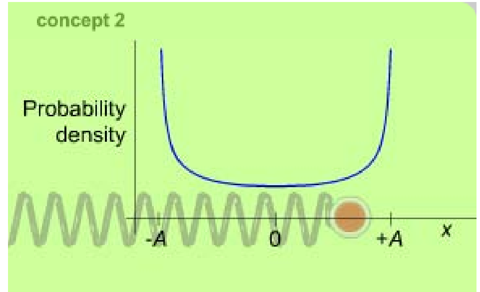


Electron fired through slits

Another conundrum can be discussed using the double-slit experiment. Electrons are fired through a pair of slits toward a barrier, as shown in Concept 1. A photographic plate records where the electrons strike the barrier. The pattern recorded on the plate (and shown in Concept 1) looks the same as that created by shining light through the slits.

Create interference pattern
Where a single electron will land can only be stated as likelihood

Although physicists can state in advance what the overall pattern will look like based on factors such as how wide the slits are, they cannot predict in advance where any given electron will land on the photographic plate. They can only state the probability that it will land near a given location − for instance, the probability is much higher in the areas of the plate that contain the most dots.

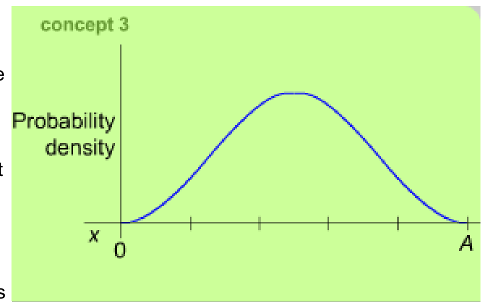


Position
Can be stated as a probability for "classical" objects

You can liken this to a game of cards. You know that on average, one out of four cards will be a spade. However, given a shuffled deck and asked to state for sure whether a spade will be the first card you draw, or the second, you cannot. You can only say that there is a one in four chance that it will be a spade.

If 12 cards are dealt from a shuffled deck of cards, the probability is greatest that 3 cards will be spades, but from observing the repetition of many such trials, you could conclude that it is fairly likely that the actual number will be 2 or 4, and it is possible that there will be 1 or 5 spades, or even 0 or 12. All you can do is describe the probability of a given number of spades that will emerge when you deal 12 cards.

The section started with a question: How can the location of a particle like an electron be described? Quantum physicists describe such a particle with a *wavefunction*. The value of a wavefunction at a particular point and time is related to the **probability** of finding the particle near that point, at that time. To be more precise, the absolute square of the wavefunction is a probability density, a concept we discuss next.

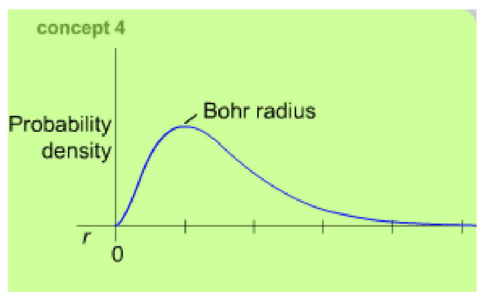


Particle in a box
Probability density describes likelihood that particle will be found in a given region

Quantum physicists use the idea of probability density to describe the likelihood of finding a particle in a given region of space. *Probability density* is the probability per unit volume and is the absolute square of the wavefunction. (In one-dimensional problems, probability density is the probability per unit length, and in two-dimensional problems, it is the probability per unit area.)

The term "absolute square" covers the case when the wavefunction has complex values (containing the imaginary number i). The absolute square equals the wavefunction multiplied by its complex conjugate. If the wavefunction contains only real values, so that its complex conjugate is the wavefunction itself, then the absolute square of the wavefunction is the same as the "ordinary" square of the wavefunction.

What is meant by probability density? It is a concept that can be applied to objects typically described with classical mechanics. For instance, someone could create a graph of your probability density at midnight. It is relatively likely that if someone is looking for you at that time, they would find you in bed. This means that the probability density function has a relatively high value there. The probability density function at midnight would also have smaller peaks at other locations in space, such as the chair in front of your computer, or in front of the refrigerator. Since the probability is zero to find you on Mars at midnight (or at any other time), the probability density function would be zero there.



Position of bound hydrogen electron
Height of graph reflects likelihood of finding particle in given region

The same idea can be applied to the position of a mass on a spring that is moving back and forth in simple harmonic motion with a certain total energy (or, physicists say that the system is in a particular state). You could graph the probability density function for the location of the mass. In fact, we did so, and that graph is shown in Concept 2.

How do you interpret this graph? Again, the locations where the graph is higher are the locations where it is more likely you would observe the mass. If you were only allowed to take photographs of the mass at random times, you would create a graph like this. (This is the situation in which quantum physicists find themselves.) The graph is higher near the endpoints of the mass's motion because it is moving more slowly there, which means it spends more time there and you are more likely to observe it there. Taking a snapshot at a random time, you are least likely to find it near the center because it moves the most quickly through that region.

In Concept 3, we move more to the realm of quantum mechanics. One intellectual construct that quantum physicists use to describe particles is the concept of a particle moving in a one-dimensional rigid box. We use the concept of waves and a wavefunction to describe that particle, which in this case is in its lowest energy state. We then construct a probability density graph. It shows that the particle is more likely to be found in the center of the box. The walls of the box are at 0 and A, and the particle never escapes the box. This means the particle is always present within the box, that is, the probability of finding it outside this region is zero.

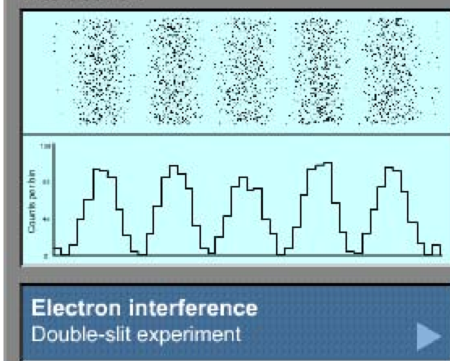
The probability density that describes the position of an actual particle is rarely as simple as the one shown in Concept 3. In Concept 4, you see the graph of the probability density function for the electron in the ground-state in a hydrogen atom (the electron has the lowest energy that it can have). You can see that the electron is most likely to be found at the Bohr radius, but that it is possible for the electron to be found at radii other than this value. For large values of the radius, the probability density approaches zero. The very low probability density at large radii means it is possible, but highly, highly unlikely, that a bound electron can be found, say, one meter from the nucleus. Quantum physicists can use the Schrödinger equation to create a wavefunction that yields the graph you see in Concept 4. This is one of the fundamental equations in quantum mechanics.

The graph in Concept 4 tells us about the behavior of an electron in the ground-state of a hydrogen atom. If you measured the distance of the electron from the nucleus, and then did so again and again, each time charting its location, you would create a graph (a *radial probability density*) similar to that. You would note that the largest number of your observations have the electron at a distance equal to the Bohr radius, but you would find it at other locations as well.

Along with Einstein's theories of relativity, quantum mechanics has changed the world's conception of reality. Einstein showed that time and length cannot be treated as absolutes, but that the motion of the observer affects these properties. Physicists have now shown that concepts as common as the location of a particle must be stated in terms of probabilities. The work of these physicists has set the direction of physics for the last 100 years or so.

42.9 - Interactive exercise: observing the probabilities of a particle

interactive 1



The simulation on the right is used to show how an interference pattern emerges as more and more electron waves pass through a pair of slits.

You are asked to observe two things.

First, press the FIRE button a few times and observe the location of the first five or six electrons. Then press RESET, and do this again. Do you observe each electron at the same location every time you run the simulation?

At the risk of ruining the punch line, the answer is no. If this is not clear, just fire one electron, press RESET, and fire one electron again.

You cannot predict in advance where a given electron will land on the screen. You can predict the probability of it landing near a certain spot, but not where it will land for sure.

Although you cannot predict the location of any given electron, after enough electrons pass through, the overall pattern is visible as predicted by the intensity function.

After you fire 200 or so electrons, look at the graph. (Depress and hold down the FIRE button.) It will very closely resemble the graph of the intensity function for the interference of two electromagnetic waves passing through a pair of slits. Where the function is greater, you will see more photons accumulate over time. The smaller the function at a point, the fewer the number of particles (be they photons or electrons) that will accumulate nearby. Try firing 500 electrons, and you will see that the graph resembles the intensity function even more as the number of particles increases.

What you are observing here is in contrast to what one would expect using classical physics. If you toss a ball with a speed of exactly 25.0 m/s at an angle of 30.0 degrees from a height of exactly 1.50 m in a vacuum, classical physics states that you can exactly predict its trajectory and where it will land. If you again throw it with the same velocity from the same height, it will always land at the same location. Classical physics would claim the same certainty and reproducibility for an electron launched in the same way.

No such luck with quantum physics − or maybe one should say "it's all luck" with quantum physics? If you fire two consecutive electrons at the double-slit system, under identical conditions, you do not know for sure that the particles will land even remotely near the same spot. You can only observe where they land this time.

The more electrons that are fired, the more accurately you can predict the overall pattern. But you can no more predict the outcome of a single spin of a roulette wheel than the landing point of any one photon or electron.

42.10 - Heisenberg uncertainty principle

equation 1

Heisenberg uncertainty principle

$$(\Delta x)(\Delta p_x) \geq \frac{\hbar}{2}$$

Δx = uncertainty in position
Δp_x = uncertainty in momentum

In prior sections, we discussed the interpretation of a matter wavefunction as relating to the likelihood that a particle will be found in a particular location.

Now we will turn to what happens when you try to measure the location of a particle. We will keep things simpler by assuming that the particle moves along a line, namely the x axis.

In general, if you want to know where a particle is, you make a measurement. All measurements are imperfect and have some uncertainty. Δx represents the uncertainty in position. If you were making a measurement of the momentum of the particle in the x-direction, there would also be some uncertainty of momentum Δp_x.

A classical physicist would agree there is some uncertainty in all measurements, just as you may have learned while doing lab assignments. Better procedures and instruments can reduce the uncertainty. For example, if you were trying to measure an electron's position and momentum 5.00 seconds after it had been propelled by a given electric field, you might set up 1000 identical trials, use the best instruments available, and average the results for momentum and for position.

According to classical physics, these uncertainties can in principle be reduced to zero. There are no limits to knowledge about a classical particle. A classical physicist considering the theoretical uncertainty of position and momentum could write an equation like:

$$(\Delta x)(\Delta p_x) = 0 \quad \text{(classical physics)}$$

$\hbar$ = Planck's constant/2π

The German physicist Werner Heisenberg, in contrast, made the bold statement that knowledge is limited. He related the uncertainties in the particle's position and momentum along the same axis:

$$(\Delta x)(\Delta p_x) \geq \frac{\hbar}{2}$$

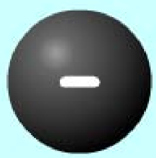
This inequality is a statement of the *Heisenberg uncertainty principle*. It states that there is a tradeoff between reducing the uncertainty in position and trying to do the same for momentum. One can find situations where Δx is relatively small, but this low uncertainty in position comes at the price of a higher uncertainty in the particle's momentum. Or if the momentum can be well determined, then the particle's location will be less precisely known.

Heisenberg's principle does not describe an "equipment problem". The problem cannot be solved with a better microscope, or more expensive lab equipment, or any other technique. Physicists hold that it is a fundamental property of reality: A particle's position and momentum must reflect a certain minimum amount of uncertainty because the particle simply cannot have both a definite position and a definite momentum. The wave/particle is spread out in space.

One important implication of this principle is that a particle can never have zero kinetic energy. If it did, it would be stationary (and have zero momentum), and then both its position and momentum could be determined. Consistency with the uncertainty relation requires that the particle must have at least some amount of kinetic energy, even at zero temperature.

Again, as with much quantum physics, the departure from classical physics becomes apparent only when the mass of the particle is very small.

42.11 - Sample problem: Heisenberg uncertainty principle



A measurement shows that an electron is located in a region between x = 0 and x = 0.2 nm. What is the minimum uncertainty in the electron's x-velocity that is consistent with the Heisenberg uncertainty relation?

Variables

uncertainty in position	Δx
uncertainty in momentum	Δp
electron mass	$m = 9.11\times10^{-31}$ kg
$h/2\pi$	$\hbar = 1.05\times10^{-34}$ J·s

What is the strategy?

1. Use the Heisenberg uncertainty principle to relate the uncertainties in position and momentum.

2. Find the uncertainty in the velocity from the uncertainty in the momentum.

Physics principles and equations

The Heisenberg uncertainty principle

$$(\Delta x)(\Delta p_x) \geq \frac{\hbar}{2}$$

Step-by-step solution

Step	Reason
1. $(\Delta x)(\Delta p_x) \geq \frac{\hbar}{2}$	uncertainty principle
2. $\Delta p_x \geq \frac{\hbar}{2\Delta x}$	solve for uncertainty in momentum
3. $\Delta p_x = m\Delta v_x$	express uncertainty in momentum in terms of velocity
4. $m\Delta v_x \geq \frac{\hbar}{2\Delta x}$ $\Delta v_x \geq \frac{\hbar}{2m\Delta x}$	solve for the velocity uncertainty
5. $\Delta x = \frac{1}{2}(0.2\ \text{nm})$	position uncertainty
6. $\Delta v_x \geq \frac{1.05 \times 10^{-34}\ \text{J} \cdot \text{s}}{2(9.11 \times 10^{-31}\ \text{kg})(1.0 \times 10^{-10}\ \text{m})}$	substitute known values
7. $\Delta v_x \geq 5.76 \times 10^5\ \text{m/s}$	evaluate

The minimum uncertainty in the velocity is very large because of the small uncertainty in the position. This example is a simple model of an electron confined to an atomic-sized region, and demonstrates the important implications of the uncertainty principle when small systems are studied.

If the uncertainty in velocity is on the order of 10^5 m/s, then typical values of the particle speed (as measured in an experiment) are also on that order. An electron moving at such a speed has kinetic energy of about an electron volt. In other words, we have just seen that Heisenberg's uncertainty principle predicts that an electron confined to an atomic-sized region will have an energy on the order of an electron volt. This is comparable to the observed ground state energy of the system, and this triumph confirms our belief (certainty?) in the uncertainty principle

42.12 - Gotchas

More massive particles always have shorter de Broglie wavelengths than less massive particles. Not necessarily. A particle's wavelength depends on its momentum, which equals its mass multiplied by its velocity. A more massive particle may have a longer wavelength than a lighter particle, if its velocity is small enough.

Only particles like protons and electrons have a de Broglie wavelength associated with them. "Matter waves" do not apply to larger things like baseballs. No. The wavelengths of objects with relatively large momenta, such as moving baseballs, are so small that experiments do not reveal their wave-like properties. But these objects are still subject to the laws of quantum physics. For large objects moving at ordinary speeds, the predictions of quantum physics and those of Newtonian mechanics are identical for all intents and purposes, much as special relativity essentially agrees with Newtonian mechanics at slow enough speeds.

42.13 - Summary

Further evidence that light has properties of particles − in addition to properties of waves − is provided by the Compton effect. When a photon and an electron collide, they behave much like two colliding billiard balls. The photon's frequency is reduced, because it has transferred some energy to the electron. The Compton effect provides evidence that though photons do not have mass, they **do** have momentum.

One of the basic tenets of quantum physics is that matter, like light, also exhibits wave-particle duality. The wave nature of matter explains why the energies of electrons in an atom are quantized.

The equation $\lambda = h/p$ relates a particle's wavelength to its momentum, and applies to particles of matter or of light.

Matter wave interference can be observed using an experiment similar to the double-slit experiment for light. However, in the case of matter waves, the wavelengths of even the smallest particles (generally those with the least momentum), such as electrons, are very small and so require very small slits to observe. Such small, closely-spaced slits are difficult to manufacture, but certain crystals naturally have atomic spacing similar to the required slit spacing. Matter waves were first observed using such crystals.

Matter waves are probabilistic. The wave- and particle-like properties of matter are reconciled in the form of the particle's wavefunction. The value of the wavefunction is related to the probability that the particle will be found at a given location at a given time. Mathematically speaking, a particle is described by a probability density function that tells the probability of observing the particle in any region.

One consequence of the wave-particle duality is that it is impossible to know both the position and momentum of a particle to infinite precision at the same time. In fact the product of the uncertainties in position and momentum has a precise lower limit: half of $\hbar$. This is the Heisenberg uncertainty principle.

Equations

Momentum of a photon

$$p = \frac{h}{\lambda}$$

Compton's equation

$$\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\theta)$$

Wavelength of a matter particle

$$\lambda = \frac{h}{p}$$

Heisenberg uncertainty principle

$$(\Delta x)(\Delta p_x) \geq \frac{\hbar}{2}$$

Chapter 42 Problems

Chapter Assumptions

1 eV = 1.60×10^{-19} J

$h = 6.63\times10^{-34}$ J·s

$\hbar = 1.05\times10^{-34}$ J·s

$m_e c^2 = 0.511$ MeV

Conceptual Problems

C.1 Electron A is accelerated from rest by a potential difference of 500 volts, and electron B undergoes a similar treatment by a potential difference of 1000 volts. Afterward, which electron has the larger de Broglie wavelength? Explain your answer.

○ Electron A ○ Electron B

C.2 An electron and a proton are each accelerated from rest by a potential difference whose magnitude is 500 volts. (The electron is accelerated from a potential of 0 V to 500 V, whereas the proton is accelerated from 500 volts to 0 volts.) Afterward, which has the larger de Broglie wavelength? Explain your answer.

○ Electron ○ Proton

C.3 A bowling ball is dropped from a tall building. While it is speeding up, does its de Broglie wavelength become smaller, become larger, or stay the same? Explain your answer.

i. Becomes smaller
ii. Becomes larger
iii. Stays the same

Section Problems

Section 1 - The Compton effect

1.1 What is the momentum of a photon of visible light with wavelength 555 nm?

_________ kg · m/s

1.2 What is the momentum of an x-ray photon with a frequency of 1.10e+18 Hz?

_________ kg · m/s

Section 2 - Derivation: the Compton effect equation

2.1 A photon with a wavelength of 7.4e−11 m is scattered by a loosely bound electron in a metallic target and exits at an angle of 135° from its original direction. What is the wavelength of the scattered photon?

_________ m

2.2 A photon emerges from radioactive cesium with an energy of 0.662 MeV. It is scattered by a loosely-bound electron in an aluminum target and a detector records its energy as 0.228 MeV. Through what angle was the photon scattered?

_________ °

2.3 Show that the units of $h/m_e c$, the Compton wavelength, are truly units of length.

2.4 A 0.31 MeV x-ray photon collides with an electron and scatters straight back the way that it came. (a) What is the energy of the backscattered photon? (b) What is the kinetic energy of the scattered electron? Assume that it was initially at rest.

(a) _________ MeV
(b) _________ MeV

2.5 In a Compton scattering event, an electron that is initially at rest is given a kinetic energy of 0.42 MeV when a photon with initial energy of 0.60 MeV strikes it. (a) What is the final energy of the photon? (b) What is the wavelength of the scattered photon? (c) Through what angle was the photon scattered?

(a) _________ MeV
(b) _________ nm
(c) _________ °

2.6 For incident photons of wavelength 5.00×10^{-11} m, what is the maximum kinetic energy that can be transferred to an electron in a Compton scattering event?

__________ eV

2.7 Incoming photons of equal energy strike electrons that are at rest. The maximum kinetic energy of a scattered electron is 61.0 keV. What is the energy of an incoming photon?

__________ keV

2.8 X-rays with a wavelength of 0.0350 nm undergo Compton scattering from a carbon target. (a) Find the wavelength of light scattered at 110° from the forward direction. (b) Find the magnitude of the momentum of a scattered electron corresponding to these scattered x-rays. Assume the electron is at rest before the collision.

(a) __________ nm
(b) __________ kg · m/s

Section 4 - Matter waves

4.1 What is the wavelength of a 1.20×10^{3} kg car that is traveling along the highway at 99.0 km/hr?

__________ m

4.2 The wavelength of a nonrelativistic electron is 2.71e−6 m. How fast is it moving? Express your answer as a fraction of the speed of light, *c*.

__________ c

4.3 The de Broglie wavelength of a particle is 7.01×10^{-13} m and it is moving at 0.189% of the speed of light. What is its mass?

__________ kg

4.4 The kinetic energy of an electron is 29.3 eV. (a) How fast is it moving, expressed as a fraction of the speed of light *c*? (b) What is its wavelength?

(a) __________ c
(b) __________ m

4.5 An electron, starting from rest, is accelerated through a potential difference and ends up with a de Broglie wavelength of 1.12×10^{-10} m. What potential difference is needed to accomplish this?

__________ V

4.6 An electron is moving such that its de Broglie wavelength equals the Compton wavelength. How fast is it moving, expressed as a fraction of the speed of light *c*? Hint: You must use a relativistic expression for momentum.

__________ c

4.7 Calculate the total energy of a proton that has a de Broglie wavelength of 0.500 femtometer (1 fm = 10^{-15} m). The rest energy of a proton (mc^2) is 938 MeV. Hint: This requires a relativistic calculation.

__________ MeV

Section 6 - Observing matter waves

6.1 A beam of neutrons, all moving at the same speed, is directed through a pair of slits whose centers are 2.00×10^{-6} m apart. An array of detectors is located 49.9 m from the slits. The intensity pattern of received neutrons is measured, and the first zero-intensity point is 1.76 mm off-axis. The mass of a neutron is 1.68×10^{-27} kg. (a) What is the de Broglie wavelength of a neutron in the beam? (b) What is the speed of the neutrons in the beam?

(a) __________ m
(b) __________ m/s

Section 10 - Heisenberg uncertainty principle

10.1 A barrier contains a vertical slit of width 5.0×10^{-6} m. An electron approaches the barrier and passes through the slit. (a) What is the uncertainty in the horizontal position of this electron (measured parallel to the barrier, not in the direction of the electron's motion) as it emerges on the other side? (b) What is the minimum uncertainty in its corresponding horizontal momentum?

(a) __________ m
(b) __________ kg · m/s

10.2 The position of a 900 kg boulder's center of mass has been determined to within an uncertainty of 1.0 nm. (a) What is the minimum uncertainty in the boulder's velocity? (b) Repeat the calculation, but for a proton with the same uncertainty in position. (c) Repeat the calculation, but for an electron with the same uncertainty in position.

(a) __________ m/s
(b) __________ m/s
(c) __________ m/s

10.3 The *y*-component of a dust particle's velocity is measured with an uncertainty of 1.0×10^{-6} m/s. The particle has a mass of 1.6e−9 kg. (a) What is the limit of the accuracy to which we can locate the particle along the *y*-axis? That is, what is the minimum uncertainty in the *y*-position? (b) Does this place any limitation on the accuracy with which we can locate its position along the *x* or *z* axes?

(a) __________ m
(b) ○ Yes ○ No

chapter 43 Nuclear Physics

43.0 - Introduction

Nuclear power plants convert the energy stored in atoms into energy humans can use.

"Turn off the lights when you leave the room." "Don't buy that gas guzzler." "Turn the thermostat down a few degrees in winter." These phrases are all about saving energy, a worthwhile topic, but this chapter is about where that energy comes from.

The energy powering your computer and lights is electrical, but that electricity was likely generated from the burning of oil or coal. In turn, these fossil fuels came from the remains of ancient animals and plants, which derived their energy from sunlight. If your power was generated from hydroelectric sources − letting water stored behind a dam spin a turbine as it falls to a lower height − it was sunlight that evaporated the water, which later turned into rainwater which was stored behind the dam.

Ultimately, the source of our energy here on Earth is the Sun, and nuclear fusion is what powers the Sun. The fusion process takes lighter chemical elements such as hydrogen, and forms helium and heavier elements, in the process releasing energy.

In that respect, we owe our very existence to nuclear energy. However, to some, the term "nuclear physics" brings darker images to mind. For better or worse, the world learned what nuclear fission was in 1945 at Hiroshima, although the exact same physics principles behind atomic bombs are used peacefully every day.

In this chapter, you will learn about the nature of atoms, which is a field of study called *nuclear physics*. You will become more familiar with protons, neutrons and alpha particles. You will learn what occurs when atoms are split apart (fission), or when they are forced to merge (fusion).

You will also learn about many peaceful applications in the field of atomic physics such as radioactive dating (no, this does not mean that two radioactives go out for dinner and a movie) and energy production in nuclear reactors.

43.1 - The atom and the electron

Nobel Laureate Richard Feynman posed a hypothetical situation in which our entire base of scientific knowledge was destroyed, and we could pass on just one sentence to our descendants. What would we choose to tell them, in order to convey the most information in the fewest words?

In the famous "Feynman Lectures", his choice was to tell the post-apocalyptic population that all matter is composed of atoms, which are tiny particles that continually move around, attracting each other when they are fairly close together, but strongly repelling when they are pressed even closer together.

How did scientists arrive at this modern picture of the atom? In fact, what is an atom?

Humans have spent many millennia asking the question "What is the world made of?" We have come a long way from the ancient Greeks, who believed that the cosmos was composed of four elements: water, earth, air, and fire. Now, we know of over 100 chemical elements, which are the fundamental building blocks of matter. You have heard of many of the elements: hydrogen, oxygen, carbon, gold, lead, and so on.

The modern definition of an *element* is a substance that cannot be divided or changed into another substance using ordinary chemical methods. Each element has different physical and chemical properties such as density, specific heat, and the way in which it bonds with other elements. You can see an example of a common element, carbon, in Concept 1.

An atom is the smallest piece of an element that still has its chemical and physical properties. Atoms consist of electrons circling a nucleus composed of protons and neutrons. This is illustrated in Concept 2. The next few sections will discuss the inner structure of the atom.

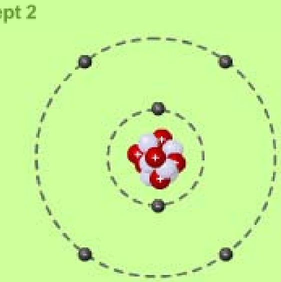
The modern theory of the atom began with the discovery of the first subatomic particle, the electron, in experiments performed over the course of about 40 years in the late 1800s. Physicists found that when they applied a high voltage across a low-pressure gas, an electric discharge was produced. They used a device called a *cathode ray tube* to achieve this effect. Today, you can witness a similar discharge in a neon sign or a fluorescent light bulb.

The physicists determined that the gas in the tube was conducting electric charge. The

concept 1

Elements
Fundamental building blocks of matter
Atoms are smallest distinct component of elements

concept 2



Atoms
Contain a nucleus and one or more electrons

tube's negative electrode, also called a *cathode*, emitted a type of "invisible ray" that could cause a glow in the treated glass wall of the tube. Later experiments showed that the rays could be deflected by electric and magnetic fields. This suggested that the rays were charged particles, and not a form of electromagnetic radiation.

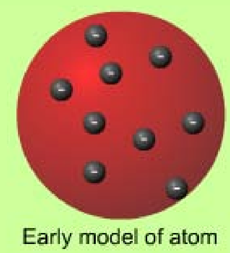
Today, scientists would say that an electric current is flowing through the gas, and the current consists of electrons. The discovery of the electron is credited to J.J. Thomson in 1897. He showed that these rays were small, negatively charged particles. Thomson also made the first measurement of the ratio of their charge to their mass.

As with many scientific discoveries, this discovery raised more questions. Scientists had long known about electric charge. Since most matter is neutral, they knew there must be positive charges to balance the negative charge of the electrons. To put it at the microscopic level: Atoms must contain positive charges to balance the negative electrons. But how were these positive and negative charges arranged in an atom? Answering that question is the topic of the next section.

43.2 - Rutherford's discovery of the nucleus

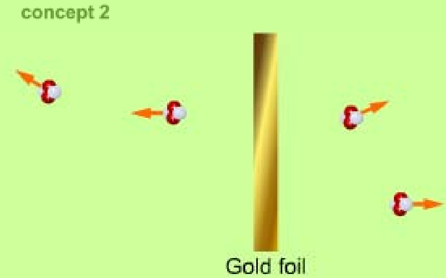
Nucleus: A relatively small region in the center of an atom where the positive charge - and most of the mass - of an atom is located.

concept 1



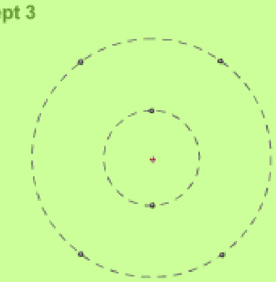
Old "plum pudding" model
Positive charge distributed evenly, with embedded electrons
· Wrong!

concept 2



Rutherford's experiment
Fired alpha particles at gold foil
Some particles scattered by a large amount
Implied compact, massive nucleus

concept 3



Rutherford's model of atom
Nucleus consists of positive charge
Nucleus is small compared to orbits of electrons

The modern model of the atom describes it as a small positive nucleus, surrounded by orbiting electrons. The radii of the orbits are far larger than the size of the nucleus. How did scientists create this model of the atom? For instance, what led them to believe that the nucleus was small compared to the size of the orbital radii?

Scientists in the early 20th century struggled to understand the nature of atoms. They could not directly see the structure of atoms, but knew that they contained negatively-charged electrons. They reasoned that atoms must contain equal amounts of positive and negative charges since matter tends to be electrically neutral.

Without other data, most scientists thought that the positive charges that make up matter were evenly distributed throughout the atom − why not? Since electrons have so little mass, scientists knew that the positive charges carried almost all the mass of the atom. Their mental picture of the atom looked like a blob of positively-charged cookie dough with small chocolate chips (electrons) embedded in it. The model actually bore the name of a more popular dessert at the time: The *plum pudding model* consisted of negatively-charged electrons (plums) scattered throughout a massive cloud of positive charge (pudding) that was distributed uniformly throughout the volume of the atom. This model is shown in Concept 1.

Lord Rutherford made a major breakthrough in this area, winning the 1908 Nobel Prize in Chemistry for his experiments. As a good scientist, he wanted data to support (or contradict) the plum pudding model of the atom. At the time, Rutherford was studying *alpha particles*, which are massive, positively charged particles that are emitted at high speed from some radioactive substances (such as radon). He realized that a beam of alpha particles might serve as a tool to probe the atomic interior.

In his experiment, he aimed a beam of these particles at a thin gold foil, and measured the distribution of the outgoing alpha particles. This is known as a *scattering experiment* − observing how the particles scatter − and it is a now-common technique to probe the details of atomic-sized systems.

A mechanical analogy to what Rutherford did would be to probe the interior of a box by shooting a high-powered BB gun at it. If the box were filled with sponge cake, the BBs would pass through. If it had a metal plate inside, the BBs would rebound back. Or if there were a small metal sphere inside, a small fraction would rebound or scatter sideways, with a distribution of angles.

Rutherford initially assumed that the alpha particles would be passing through a "pudding" of positive charges spread uniformly throughout the foil. Relying on this model, Rutherford predicted that most of the alpha particles should just pass straight through or be only minimally deflected.

This was not what he observed. Much to Rutherford's surprise, a small fraction of the particles were scattered by 90° or more. Occasionally, an alpha particle even rebounded from the foil, straight back at the source. He concluded that the atom must not have a uniform distribution of positive charge inside. Instead, the large force necessary to cause such scattering of the positively charged alpha particle could be provided only if the atom's positive charge (and mass) were highly concentrated within the atom, in a region called the nucleus.

After analyzing the data and seeing how few of the alpha particles actually scattered, Rutherford concluded that the nuclear radius must be about 10,000 times smaller than the atomic radius, a figure that is still accepted today. Rutherford's groundbreaking experiment proved that the atom is mostly empty space. His atomic model is shown in Concept 3. Note that the diagram is not even close to being drawn to scale; the atomic diameter is far too small compared to the nuclear size. If the diagram were drawn so that the nucleus were 1 cm wide, roughly the width of your pinky, the atom would have to be drawn about 100 meters wide, about 10% longer than the length of an American football field.

43.3 - Components of the nucleus

Rutherford proved that atoms consist of a compact, very dense positively-charged nucleus surrounded by negatively-charged electrons. In this section, we take a deeper look at the parts of the nucleus, and introduce some common terminology and notation that scientists use when talking about elements and nuclei.

The positive charge of the nucleus comes from particles called protons, which are about 1800 times more massive than electrons. The simplest nucleus consists of a single proton, and the simplest atom is hydrogen, which consists of a proton and an orbiting electron.

What distinguishes an atom of hydrogen from an atom of gold? A chemical element is defined by the number of protons in its nucleus. For an atom of a particular element, the nucleus consists of Z protons, where Z is called the *atomic number* of the element.

Hydrogen has a single proton, so for hydrogen, $Z = 1$. Gold has 79 protons, so its atomic number is 79. Protons and electrons have equal but opposite charges. This means that an electrically neutral atom has the same number of electrons as protons, so for example a gold atom has 79 protons and 79 electrons.

There can be more to a nucleus than just protons; there may also be other particles present, called *neutrons*. These are uncharged particles that are just a bit more massive as the proton (about 0.1% more massive). Protons and neutrons are known as *nucleons* because they make up the nucleus.

The *neutron number*, N, states the number of neutrons. Any atom of an element always has the same number of protons, but it can have different numbers of neutrons. Two forms of an element with different numbers of neutrons are known as *isotopes*. Hydrogen, for instance, **always** has a single proton, but it can have either zero neutrons ("common hydrogen"), one neutron (an isotope called deuterium), or two neutrons (called tritium). Deuterium and tritium are shown in Concept 2.

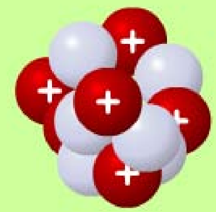
Different isotopes of an element will have different atomic masses because of the differing numbers of neutrons. The table in Concept 3 summarizes the charge and mass properties for protons, neutrons, and electrons.

Atomic masses can be measured using instruments such as the mass spectrometer. Masses are commonly given in terms of atomic mass units, u, defined such that the mass of the most abundant kind of carbon atom, carbon-12, has a mass of exactly 12 u. The value of an atomic mass unit in kilograms is given in Equation 1.

The sum of Z, the number of protons, and N, the number of neutrons, is called A, the *mass number* of the atom. Since electrons have little mass compared to protons and neutrons, the mass number is very close to the entire mass of the atom when it is expressed in atomic mass units.

An atom with a particular combination of Z and N is called a *nuclide*. A particular nuclide always has the same type of nucleus. A widely adopted notation to identify a nuclide is to write the chemical symbol of the element, with its atomic number subscripted to the left and its mass number superscripted to the left. This is shown in Equation 2, as applied to the most common carbon nuclide which has six protons and six neutrons.

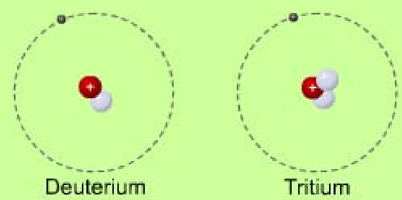
concept 1



Parts of the nucleus

Protons
· Z = number of protons
Neutrons
· N = number of neutrons

concept 2



Isotopes of hydrogen

Isotopes

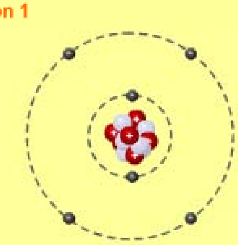
Same number of protons, different number of neutrons

concept 3

	Mass	Charge
Proton	1.673×10^{-27} kg	$+1.602 \times 10^{-19}$ C
Neutron	1.675×10^{-27} kg	0 C
Electron	9.109×10^{-31} kg	-1.602×10^{-19} C

Neutrons and protons have similar mass
Protons far more massive than electrons

equation 1



Carbon-12: mass = exactly 12 u

Atomic mass unit

$u = 1.66054 \times 10^{-27}$ kg

u = atomic mass unit (amu)

equation 2

Carbon-12 has 6 protons and 6 neutrons

$^{12}_{6}C$

Atomic number = 6
Mass number = 12
Carbon-12: mass = exactly 12 u

Mass number

$A = Z + N$

A = mass number
Z = atomic number
N = neutron number

43.4 - Interactive checkpoint: using Z, N and A

A neutral atom of sodium-23 has 11 protons and a mass number of 23. How many electrons does it have? How many neutrons does it have?

Answer:

Number of electrons =

Number of neutrons =

43.5 - The strong nuclear force

Atoms contain positively-charged nuclei that attract the negative electrons, and the nuclei contain closely packed protons and neutrons. These conclusions bring up a major question about the nucleus. Since protons are positively charged, and positive charges repel, why do the protons in the nucleus not fly away from one another? In this section, we will explore the nature of the force that holds nucleons together.

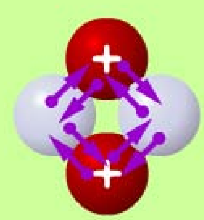
One good hypothesis would be that the force of gravity attracts them. This is a good hypothesis, but incorrect: It turns out that gravity is far too weak. (For two protons, by using Newton's law of gravity and Coulomb's law, you can calculate that the attractive force of gravity is about 10^{36} times weaker than the electrostatic repulsive force.)

One hypothesis down. Since gravity cannot explain the stability of the nucleus, the only alternative is that there must be another attractive force. This fundamental force is called the *strong force*.

The strong force has several important properties. It is "strong;" it manages to hold together the protons in a nucleus despite their electrostatic repulsion. It always binds particles together, even if their electric charges are the same, or if they are uncharged.

The strong force causes protons to attract protons, protons to attract neutrons, and

concept 1



Strong force
Holds particles in nucleus together
Is very strong!
Always attractive, regardless of charge
Acts only over a very short range

neutrons to attract other neutrons. The strong force acts only over a very short range.
For example, once two protons are separated by more than about 10^{-15} m (roughly their own diameter), there is hardly any attraction due to the strong force, though the electrostatic repulsion is still substantial.

Although much has been learned about the properties of the strong force from experiments, there is no simple formula to relate its strength to distance. With the electrostatic and gravitational forces, the amount of force is inversely proportional to the square of the distance between the particles. In contrast, no simple formula can be stated for the dependence of the strong force on distance, though there are complicated numerical approximations.

How have physicists studied the strong force? They experiment by bombarding target nuclei with high energy particles, which are influenced by the nuclei via the strong force or the electrostatic force. These forces can change the paths of the incoming particles. By observing the distribution of the outgoing particles, and comparing it to the predictions of theoretical models, scientists can test these models.

43.6 - Nuclear properties

Are the nucleons rigid objects, or soft, compressible ones? In other words, do they behave like hard marbles that are clumped together, or is there some flexibility to them, like cotton balls being crammed into a bag? These questions can be answered if we can determine how the size of the nucleus depends on the number of nucleons in the nucleus.

It turns out that nucleons are nearly incompressible. This conclusion can be drawn by looking at how the radius of the nucleus relates to the number of nucleons inside. The same experiments that physicists perform to study the strong force, where they fire particles at nuclei, have also allowed them to measure other properties of the nucleus, such as its size.

They have determined that the radius of an atom's nucleus is proportional to the cube root of its mass number. As shown in Equation 1, the radius equals the cube root of the number of neutrons and protons, multiplied by 1.2×10^{-15} m.

There are some striking implications of this simple-looking formula. It holds the answer to our question about the rigidity of nucleons, and implies other facts about the nucleus. Consider how the radius of a sphere relates to its volume. The volume of the spherical nucleus is equal to $4\pi/3$ times the radius cubed. If you cube the radius, using the equation to the right, you are cubing $A^{1/3}$, which equals A. In other words, the volume is proportional to A, the number of neutrons and protons: Each time you add a nucleon, you are adding roughly the same amount of volume to the nucleus.

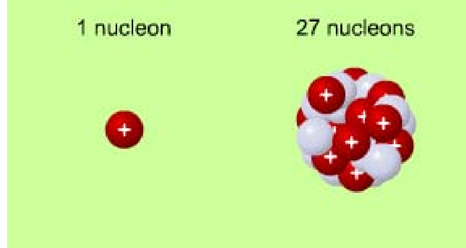
The equation also allows one to conclude that the neutrons and protons must be tightly packed. If there were large spaces between nucleons, then as their number increased, the volume would increase at an even faster rate; for example, the volume would more than double when you doubled A. (If this is not obvious to you, consider the change in the volume of adding a tenth planet beyond Pluto, versus adding another marble to a bag of marbles. Adding another planet would increase the volume of the Solar System by more than the volume of the planet itself; but as you add hard marbles to a cluster of marbles, the volume of the cluster increases by about the volume of a single marble each time.)

The equation tells scientists that the density of all nuclear material is constant. How do we know this? Density equals mass divided by volume, and since the mass and volume increase at the ratio of 1:1, the density does not change.

concept 1

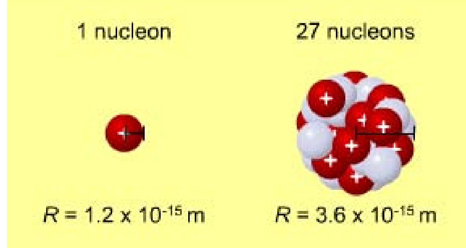
Nuclear properties
Nucleons are nearly incompressible, tightly packed
Nuclear density is roughly the same for all atoms

concept 2



Nuclear radius increases with mass number, *A*

equation 1

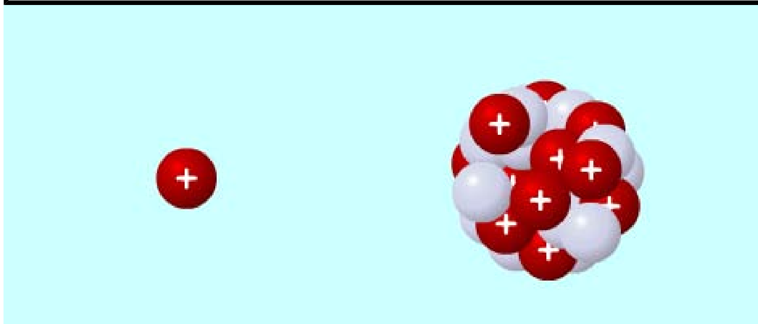


Dependence of radius on *A*

$$R = (1.2\times10^{-15}\text{ m})A^{1/3}$$

R = nuclear radius
A = mass number

43.7 - Sample problem: nuclear density



Calculate the density of the hydrogen nucleus (A = 1, mass ≈ 1.0 u) and of an aluminum-27 nucleus (A = 27, mass ≈ 27 u). Express the answer in kg/m^3, to two significant figures.

Variables

mass of hydrogen nucleus	m_H
radius of hydrogen nucleus	R_H
volume of hydrogen nucleus	V_H
mass of aluminum nucleus	m_{Al}
radius of aluminum nucleus	R_{Al}
volume of aluminum nucleus	V_{Al}

What is the strategy?

1. Convert the mass of each nucleus from atomic units to kilograms using the conversion factor stated below.
2. Find the radius of each nucleus using the relationship between radius and mass number.
3. Calculate the volume of each nucleus using the radius just calculated.
4. Divide mass by volume to find the density.

Physics principles and equations

The nuclear radius grows as the cube root of the mass number.

$$R = (1.2\times10^{-15}\text{ m})A^{1/3}$$

The nuclear shape may be modeled as a sphere. The volume of a sphere in terms of its radius is

$$V = \frac{4\pi}{3}R^3$$

The definition of an atomic mass unit is

u = 1.66 × 10^{-27} kg

Mass density

$$\rho = m/V$$

Step-by-step solution

We begin by converting the mass of the hydrogen nucleus into kilograms, then we calculate the radius of the nucleus. Using the radius, we calculate the nuclear volume. Finally, we divide the mass by the volume to determine the density.

Step	Reason
1. $m_H = 1u = 1.66\times10^{-27}$ kg	hydrogen nuclear mass, definition of u
2. $R = (1.2\times10^{-15}\text{ m})A^{1/3}$ $R_H = (1.2\times10^{-15}\text{ m})(1)^{1/3}$ $R_H = 1.2\times10^{-15}$ m	apply nuclear radius equation
3. $V_H = \frac{4\pi}{3}R_H^3$ $V_H = \frac{4\pi}{3}(1.2\times10^{-15}\text{ m})^3$ $V_H = 7.2\times10^{-45}\text{ m}^3$	calculate volume of hydrogen nucleus
4. $\rho_H = \frac{m_H}{V_H}$ $\rho_H = \frac{1.66\times10^{-27}\text{ kg}}{7.2\times10^{-45}\text{ m}^3}$ $\rho_H = 2.3\times10^{17}\text{ kg/m}^3$	definition of density

Now perform the same calculations for aluminum.

Step	Reason
5. $m_{Al} = 27\text{ u} = 27(1.66\times10^{-27}\text{ kg})$ $m_{Al} = 4.48\times10^{-26}$ kg	aluminum nuclear mass, definition of u
6. $R = (1.2\times10^{-15}\text{ m})A^{1/3}$ $R_{Al} = (1.2\times10^{-15}\text{ m})(27)^{1/3}$ $R_{Al} = 3.6\times10^{-15}$ m	apply nuclear radius equation
7. $V_{Al} = \frac{4\pi}{3}R_{Al}^3$ $V_{Al} = \frac{4\pi}{3}(3.6\times10^{-15}\text{ m})^3$ $V_{Al} = 2.0\times10^{-43}\text{ m}^3$	calculate volume of aluminum nucleus
8. $\rho_{Al} = \frac{m_{Al}}{V_{Al}}$ $\rho_{Al} = \frac{4.48\times10^{-26}\text{ kg}}{2.0\times10^{-43}\text{ m}^3}$ $\rho_{Al} = 2.2\times10^{17}\text{ kg/m}^3$	definition of density

The densities are nearly identical. This is further confirmation that nucleons are tightly packed and incompressible.

The nuclear density is far beyond the density of materials in our experience. For example, consider gold, which is 19.3 times denser than water and almost 1.7 times as dense as lead. A nucleus is 12,000,000,000,000 times denser than gold. (In case you were wondering, no, the zero key did not get stuck down while we typed that number.) Recall that Rutherford found that the atomic radius was on the order of 10,000 times the nuclear radius. The volume of a sphere scales as the cube of the radius, so the ratio of atomic to nuclear volume is approximately $(10{,}000)^3$. Since the mass of an atom mostly resides in the nucleus, and so much of an atom is empty, this explains the incredible density of a nucleus.

43.8 - Stable nuclei

What are the rules for the number of allowed protons and neutrons in a nucleus? Is any combination of protons and neutrons possible? Could there be a hydrogen atom whose nucleus has 1 proton and 7 neutrons? What about a silver atom with 47 protons and no neutrons? Could there be an element with 150 protons and any number of neutrons?

The nuclei described above do not exist. You could not even momentarily create nuclei with such extreme imbalances of protons and neutrons, or in the last case, a nucleus with so many protons. Only certain combinations of protons and neutrons can form nuclei, and even fewer combinations can form stable nuclei.

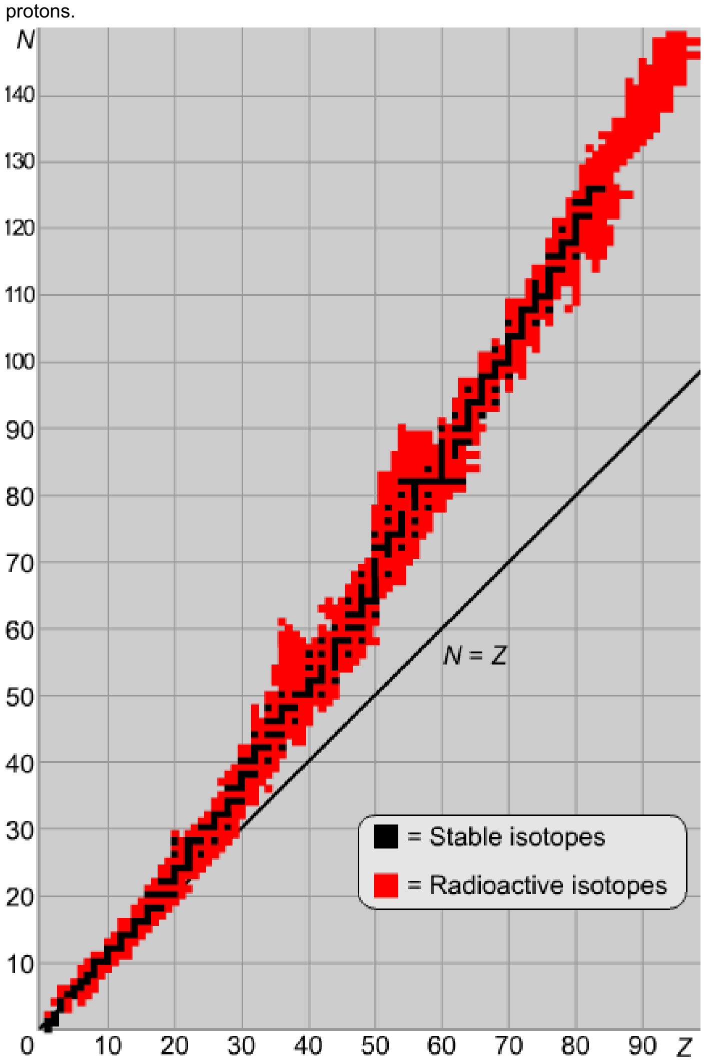
By *stable*, we mean elements that will not spontaneously decay. Gold-197 is the only stable isotope of gold, while there are several stable oxygen isotopes: oxygen-16, oxygen-17 and oxygen-18.

Some nuclides are *unstable*, meaning they have a limited lifespan. An unstable nucleus is called a *radionuclide*, and it will spontaneously and rapidly *decay* or split up into more stable pieces. Such materials are *radioactive*, and the details of the decay process are the subject of another section.

In this section, we discuss what makes for a stable nucleus. The question of stability is of fundamental importance. If all elements were unstable, then life would as we know it would not exist − the carbon, oxygen and other elements that make up your body would be constantly changing into other elements. Life is complicated enough!

On the other hand, if all elements were equally stable, then the nuclear fusion process that powers stars (including the Sun) would never happen, and the universe would be almost all hydrogen, with none of the heavier elements that make life possible.

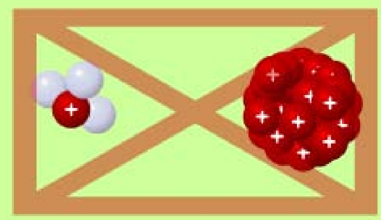
To explain why some atoms are stable and others are not, it helps to consider a diagram of stable and unstable nuclides where Z is plotted against N. For example, even though isotopes of silver with its 47 protons ($Z = 47$) have been created with mass numbers as low as 96 and as high as 124, just two of these nuclides are stable ($A = 107$ and 109). In other words, there must be 60 or 62 neutrons along with the 47 protons.



concept 1

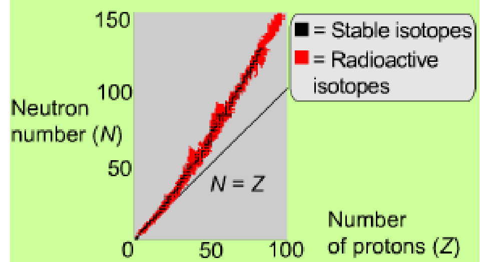
Stable nuclei
Stable only for certain combinations of neutrons and protons

concept 2



Unstable nuclei
Other combinations do not exist or are unstable

concept 3



Graph of stable and unstable nuclei
Values clustered near "band of stability"
Ratio of neutrons to protons increases with nuclear size
Larger nuclei: more neutrons required to dilute protons

We would like you to observe three important features of this diagram. First, note that the stable nuclides are clustered around a band running through the diagram. The unstable nuclides exist on either side of this band. Second, you can see how the stable nuclei are distributed. Roughly speaking, for less massive nuclei, the number of protons and neutrons is approximately equal: They cluster around the line $N = Z$. In contrast, for more massive nuclei, the number of neutrons exceeds the number of protons, $N > Z$. Third, observe that there are no stable nuclei beyond bismuth ($Z = 83$).

Since neutrons are effective at diluting the repulsive electric force between protons (by spacing them out more), and the strong force binds neutrons effectively to protons and other neutrons, it seems like having more neutrons can only bind the nucleus more tightly. You may be wondering why an atom cannot have an extremely high ratio of neutrons to protons. For instance, why are there no hydrogen isotopes with eight neutrons, or even just seven neutrons? Quantum mechanical principles dictate why.

More advanced models of the nucleus state that there are nuclear energy levels that nucleons occupy, very much analogous to atomic energy levels that electrons occupy. It turns out that neutrons (and protons) obey an exclusion principle that is analogous to the Pauli exclusion principle for electrons, meaning that two or more neutrons cannot have the same quantum-mechanical state.

Not all the neutrons will fit into one energy level. Adding more neutrons means that the additional ones must occupy higher nuclear energy levels, which raises the overall nuclear energy and reduces its stability. In such a situation, the nucleus can lower its overall energy dramatically by ejecting a neutron in a higher energy level, if the overall benefit to the nucleus is greater. In other words, there is an energy penalty to having too many neutrons around, even though neutrons do not repel each other electrostatically as protons do. This is a purely quantum-mechanical effect.

In contrast, explaining why very large nuclei are unstable, even if N is closer to Z, only requires considering the nature of the strong and electrostatic forces, not quantum mechanics.

Heavy elements like uranium have a large number of protons, which all repel one another. As the number of protons increases, the repulsive force keeps growing and growing, making the nucleus more and more unstable.

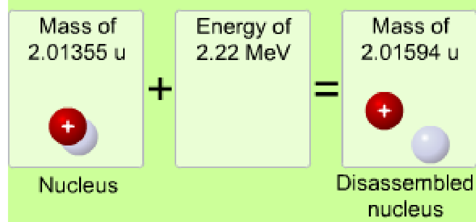
Neutrons counteract this growing instability by increasing the distance between protons, which decreases the electrostatic forces, and by attracting each other and the protons with the strong nuclear force.

However, there comes a point when the nucleus gets too large, and will be unstable no matter how many neutrons are present. This can be understood by considering the relative ranges of the strong and electrostatic forces. The strong force only acts between very close neighbors, while the repulsive electrostatic force acts between all protons regardless of their position within a nucleus. When there are lots of protons already present, and one more proton is added, it will be subject to repulsive electrostatic forces from every proton that is already there, while the attractive strong force will only be exerted by a few very close neutrons or protons. Eventually, it becomes impossible for the nucleus to "hold in" an additional proton because the strong force cannot overcome the electrostatic force.

43.9 - Nuclear binding energy

Binding energy: The energy that must be added to disassemble, or unbind, a nucleus into the protons and neutrons that make it up.

concept 1



Binding energy
Energy that must be added to disassemble nucleus completely
Increased energy of separate particles reflected in increased mass

concept 2

$$E = mc^2$$

Binding energy becomes mass
Separate particles have more mass than assembled nucleus

You may have heard of radioactivity, and know that uranium atoms will spontaneously decay into other elements, while other elements such as common iron (iron-56) are stable. Stable nuclei such as iron all have one thing in common: Their nucleons are tightly bound. Unstable atoms are not as tightly bound. What does it mean to be "tightly bound"?

Note that the term "stability" in nuclear physics is not making a statement about the tendency of an atom to enter into chemical reactions. For example, we say that iron is "stable" in the nuclear sense, even though it rusts. When iron combines with oxygen to form iron oxide, it is a chemical reaction, not a nuclear reaction. The iron remains iron when it becomes iron oxide; it shares electrons with oxygen but the element's nucleus remains unchanged.

What makes the protons and neutrons in a radioactive uranium atom less tightly bound than the nucleons in an extremely stable iron atom? Can we quantify and compare the stability of nuclei?

As it turns out, we can. One way to measure the stability of a nucleus is to try and rip the nucleons apart, overcoming the strong force. When physicists conduct such experiments, they find it takes energy to break the nuclear bonds (which the strong force is responsible for), that is, to take apart a nucleus into separate protons and neutrons. (This makes sense, because if no energy was required to separate them, they would fall apart on their own.)

The particles that emerge when the nucleus is forced apart can be analyzed. Careful measurements show that the sum of the masses of the separate nucleons is always *greater* than the mass of the nucleus when it is whole.

Concept 1 shows this, using an isotope of hydrogen, deuterium, as an example.

Why should the mass of the nucleus **increase** when it is broken up? To a classical physicist, unacquainted with Einstein's theory of special relativity, this would be a surprise since mass is assumed to be conserved.

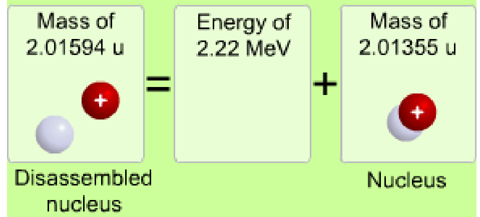
However, we know that another conservation principle applies here: the total of mass and energy (or mass-energy) must remain the same, though the individual terms may vary. Einstein's principle of mass-energy equivalence, summed up by the equation $E = mc^2$, applies. This is shown in Concept 2. It takes energy to separate the particles, and the energy added to the nucleus to fragment it into nucleons shows up as the "extra" mass. (We will assume for the sake of simplicity that the kinetic energies of the particles and of the nucleus are negligible.)

The energy that must be **added** to completely disassemble the nucleus is known as the *binding energy*. This works in both directions. The binding energy is **released** when the protons and neutrons come together to form a bound nucleus. This is illustrated in Concept 3.

The terminology could be a little confusing. You can think of it like gravity: You must "add" energy to pull apart two particles, or lift a rock farther from the surface of the Earth. That is analogous to the binding energy.

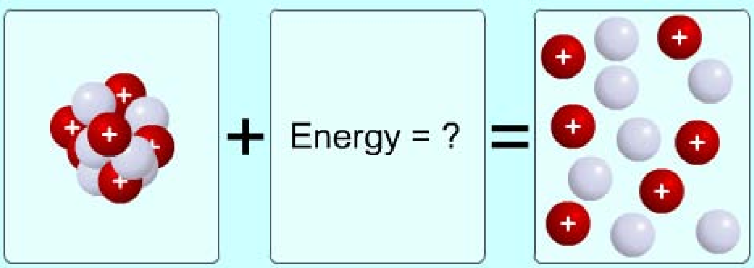
Because of the equivalence of energy and mass, the binding energy may also be related to mass. When energy is added to a nucleus to disassemble it, the mass of the parts increases. On the other hand, when a nucleus is assembled, the **release** of binding energy from the system shows itself as a corresponding **reduction** of mass in the assembled nucleus. This is sometimes called the "missing mass".

concept 3



Assembling the nucleus
Binding energy released when nucleus is assembled
Decreased energy of nucleus reflected in decreased mass
Mass becomes binding energy

43.10 - Interactive checkpoint: calculating the binding energy



For a carbon-12 nucleus with 6 protons and 6 neutrons, and a mass of 11.9967 u, how much more mass do the individual nucleons have than the assembled carbon nucleus? This is called the mass excess. What is the binding energy of carbon-12's nucleus?
A neutron has a mass of 1.0087 u and a proton has a mass of 1.0073 u.
1 u = 1.66×10^{-27} kg.

Answer:

mass excess = ______ u

binding energy = ______ J

43.11 - Binding energy curve

The binding energy is a measure of stability for a nucleus, since the binding energy is how much energy it takes to completely disassemble a nucleus. Roughly speaking, the higher the binding energy, the harder it is to pull all the nucleons apart.

However, determining which nuclides are stable is not as simple as calculating the binding energy. In this section, we discuss how the stability of a nuclide can be determined.

We will use two nuclides as examples: $^{56}_{26}\mathrm{Fe}$ (iron-56), and $^{235}_{92}\mathrm{U}$ (uranium-235). Using a table of nuclear data, one may find that the binding energy of an iron-56 atom, which has 26 protons and 30 neutrons, is 492 MeV. The binding energy of a uranium-235 atom $(Z = 92, N = 143)$ is 1784 MeV.

Does this mean that the uranium nucleus is more stable since it has a greater binding energy? Not necessarily. To compare the stability of different nuclei, a useful number to consider is the binding energy **per nucleon**. To calculate this ratio, divide the binding energy of the nucleus by the mass number, the total number of protons and neutrons. This provides a metric to compare the binding energy per nucleon in the iron isotope with the binding energy per nucleon in uranium.

concept 1

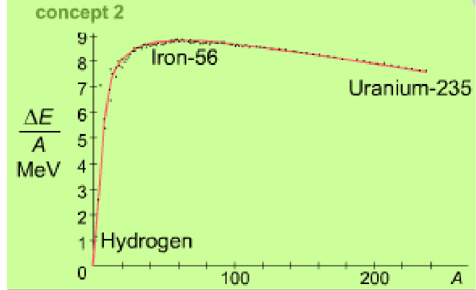
	Total binding energy	Mass number, A	Binding energy per nucleon
Uranium-235	1784 MeV	235	7.59 MeV
Iron-56	492 MeV	56	8.79 MeV

Stability and binding energy
Stability determined by binding energy per nucleon

The binding energy per nucleon in uranium-235 is 7.59 MeV and the binding energy per nucleon in iron-56 is 8.79 MeV. The binding energy per nucleon is a good measure for stability; the fact that the binder energy per nucleon is higher for iron than uranium correctly predicts that iron-56 is more stable than uranium-235. (In fact, the binding energy per nucleon for iron-56 is among the highest for all nuclides.)

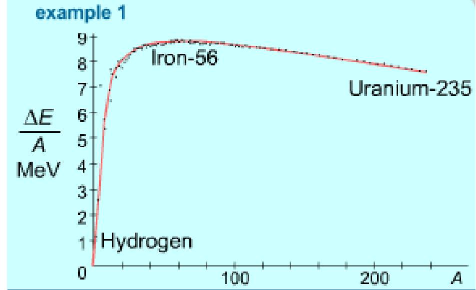
The graph in Concept 2 shows the binding energy per nucleon for naturally occurring isotopes, plotted against the mass number A. This is called the binding energy curve.

concept 2



Graph of binding energy per nucleon versus mass number
Graph peaks near iron-56

example 1



What is the total binding energy for a uranium-235 nucleus?

$$\text{binding energy} = \left(\frac{\Delta E}{A}\right) \times (A)$$

$$\text{binding energy} = (7.59 \text{ MeV}) \times (235)$$

$$\text{binding energy} = 1784 \text{ MeV}$$

43.12 - Shape of the binding energy curve

The graph of binding energy per nucleon versus mass number has a distinct shape that proves to be very important. The higher on the graph an element is (indicating more binding energy per nucleon), the more stable it is.

The most stable elements are at the highest points, with iron-56 in this region, as you can see in Concept 1.

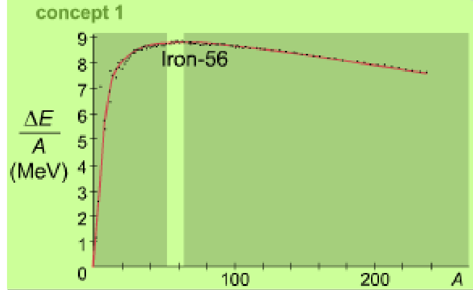
Very light nuclei (on the left of iron-56 in the binding energy curve) can become more stable if they combine to form larger nuclei through a process called fusion. By this process, the binding energy per nucleon is raised, which means that energy is released. This is the process by which stars, like the Sun, continually transform their mass into energy. In a multistep process within the Sun, hydrogen nuclei fuse together to become helium-4 nuclei.

As mentioned, the most stable locations on the curve represent elements such as iron and nickel. Heavier, radioactive nuclei to their right can increase their binding energy per nucleon and become more stable by “moving to the left and up” on the curve. For instance, you can see in Concept 3 that uranium is less stable than iron. A heavy nucleus could become more stable by emitting particles and becoming slightly smaller (the process of radioactive decay) or, in extreme cases, by splitting into two medium-sized nuclei (a process called fission). This is the principle behind radioactivity and nuclear power.

The shape of the graph also illustrates the relative distances at which the strong and electrostatic forces effectively operate. The argument that follows is reasonably complex but provides a good example of how graphical data can be analyzed. The difference between these two forces can be used to explain why the graph first shows a rapid increase of binding energy per nucleon, then levels off, and finally declines.

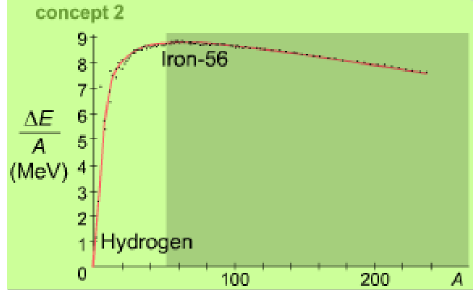
Earlier, we discussed the short-range nature of the strong force. It is so short-range that it acts only between a nucleon and its nearest neighbors. The graph supports this hypothesis. Why? When there are only a few nucleons, they are all very close and every nucleon interacts with every other. For instance, when there are two nucleons,

concept 1



Binding energy per nucleon versus mass number
Highest in the middle

concept 2



Lighter elements can undergo

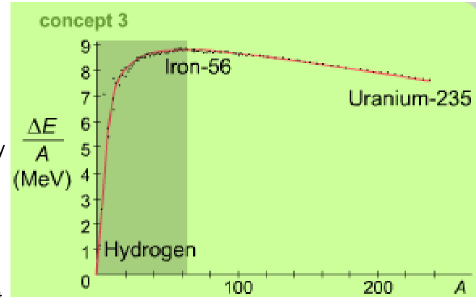
they are next to each other, and exert a strong force on one another. As a third nucleon is added, it has two neighbors to exert a force on, so the force increases faster than the number of nucleons. This means the binding energy per nucleon increases, so the line has a positive slope. (Mathematically, the binding energy of the smaller nuclei increases as the square of the number of nucleons.)

As more nucleons are added, at some point they are too far apart to all be "neighbors". For instance, when there are 100 nucleons, and another is added, it can only interact with its close neighbors. A nucleon on one side of the nucleus is too far away to exert a significant strong force on one on the far side. Adding a nucleon does not increase the binding energy per nucleon. This means with larger nuclei, the additional binding energy per nucleon becomes constant. The sharp increase in binding energy per nucleon ceases.

The strong force needs to be contrasted with the electrostatic force, which acts to push apart the protons, and which acts at a greater distance than the strong force. When a new proton is added, it is attracted only to its nearest neighbors via the strong force, but is repelled by every other proton that is already present because the electrostatic force acts at a greater range. This, in turn, makes it easier to disassemble the nucleus when more of the protons want to be separated.

In sum, at first the strong force dominates, causing the increase in binding energy per nucleon. But as the nucleus grows in size, the electrostatic force plays a larger role, causing an eventual decrease in binding energy per nucleon.

fusion and release energy



Heavier elements can divide and release energy

43.13 - Fission

Fission: A heavy nucleus breaks up into two smaller ones, releasing energy.

You may have heard the term "splitting the atom" as something that humans first accomplished in the 20th century. In this section, you will learn what it means to "split" an atom. Fission is the process used both in *nuclear reactors* to produce electrical power and also in the first *atomic bombs*.

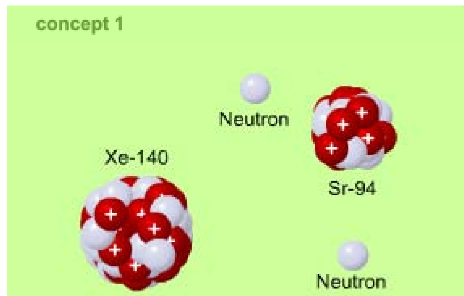
When a nucleus breaks up into smaller, more stable pieces, this is known as *fission*. Some unstable nuclei, such as uranium-236, do this spontaneously. When an atom undergoes fission, it changes identity, as the new nuclei it breaks into have different numbers of protons. There are many ways that the nucleus can break up. For example, the uranium-236 nucleus can break into Xe-140 and Sr-94, in the process releasing two neutrons. To see this fission process, press the refresh button in your browser and look at Concept 1.

However, being the impatient race that we are, humans learned to induce the process to happen at a greater rate. Induced fission, also known informally as "splitting the atom," was first performed by Otto Hahn and Fritz Strassmann. They bombarded uranium with neutrons and found that lighter elements (such as barium) were produced.

Why are neutrons effective at inducing fission? Was it not stated earlier that the strong force they supply is crucial to holding a nucleus together? Yes, but it is possible to have too much of a good thing. For a given number of protons in a nucleus − for a certain element − the band of stability is quite narrow. Too few neutrons or too many neutrons make the nucleus unstable.

In fact, neutrons are a natural choice to induce fission. Since they are electrically neutral, they can easily approach and hit the nucleus without being repelled by electrostatic forces. If the incoming neutron has the right speed, the nucleus captures it and becomes even more unstable. The nucleus has one neutron too many for the number of protons, and that immediately causes the new compound nucleus to fission into various elements.

Energy is released during fission; in the above process, roughly 200 MeV. You can understand why this happens by returning to the concept of binding energy per nucleon, which is lower for very heavy nuclei than it is for intermediate-size nuclei.



Fission
Nucleus breaks up into smaller elements
· And releases neutrons
· And releases energy

Fission and chain reactions
Neutrons can be used to cause fission
Fission can be self-sustaining
Releases energy

The binding energy is the amount of energy that is required to separate a nucleus into its separate nucleons, or it is the energy **released** when separate nucleons are brought together to form a nucleus. The more tightly bound a nucleus is, the higher the binding energy, that is, the more energy is released. Since the intermediate-size nuclei have the highest binding energy per nucleon, this means that heavier nuclei will release energy as they split apart and become medium-sized.

We have discussed fissioning a single atom with the release of energy. This is a crucial scientific achievement, but in order for this to be useful (for example, in a nuclear reactor), the process needs to be self-sustaining.

What makes fission a practical process is that in a fission reaction, one or more neutrons may be released, which can then induce more fission reactions in nearby atoms, which produce more neutrons, and so on. Why are there "extra" neutrons? Recall that for heavier nuclei, the number of neutrons exceeds the number of protons, while for lighter nuclei, the number of neutrons and protons tends to be nearly equal. This means there are usually some neutrons left over after the nuclear re-arrangement.

When there are enough uranium atoms so that at least one neutron, on average, is captured by another uranium atom, the *critical mass* has been reached. The process is self-sustaining, and it is called a *chain reaction*. Press the refresh button to see the fission process occur in Concept 2.

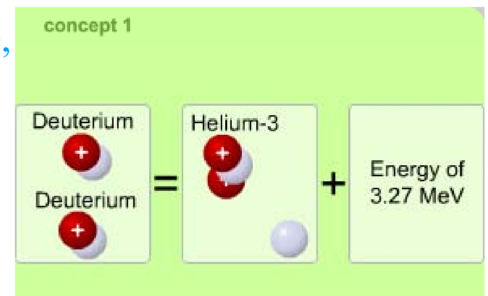
Since humans learned to split the atom in the 20$^{\text{th}}$ century, the process has been put to great use. The most well-known application was the use of nuclear fission in the "atomic" bomb. A runaway chain reaction happens very fast, releasing a lot of energy in a burst which can be used to devastating effect.

On the positive side of the ledger, nuclear power has been used as an energy source. A slowly progressing chain reaction produces a steady flow of heat that can be used to boil water, which then drives steam turbines to generate electricity. The poster-child for nuclear power is France, which supplies about three-quarters of its electrical needs with nuclear reactors.

However, nuclear power is not without its risks or costs. Some of the byproducts of the fission process are highly radioactive and remain dangerous for tens of thousands of years. A typical way to dispose of these wastes is to bury them deep in the Earth. If they leak, they contaminate water sources.

43.14 - Fusion

Fusion: Two light nuclei fuse into a heavier one, releasing energy.



Fusion
Light nuclei fuse into larger, more stable pieces
Releases energy

Fusion is a process in which nuclei join together to become a single, larger nucleus. This process also releases energy. Because positively charged nuclei repel one another, fusion does not occur spontaneously under normal conditions on Earth.

However, fusion is commonplace in the Sun and other stars where hydrogen atoms fuse into helium atoms. Fusion provides the energy to keep the star going, which Earth ultimately experiences as light and heat. Fusion occurs in the Sun because of the high temperature within the star; its interior is at about 100 million Kelvin.

At this high temperature, the atoms are in an ionized state of matter called *plasma*. While normal matter consists of distinct neutral atoms, plasma is a "soup" of positive nuclei and negative electrons. In the interior of the Sun, nuclei are hot enough and moving quickly enough to overcome the electrostatic repulsion of their positive nuclei. They move close enough to be bound by the attractive strong force between them, then fuse into a single nucleus.

Why is energy released in the process? For light elements (atoms to the left of the peak on the binding energy curve), the binding energy per nucleon increases with atomic number. As smaller nuclei are fused into larger ones, the result is a more efficient arrangement of nucleons; one that is harder to break apart. Since the new nucleus is more efficient − more tightly bound − than the previous ones, there is energy to spare.

It may be tempting to think that fusion and fission are opposite processes, since one combines nuclei and the other splits them. It may seem confusing as to how two "opposite" processes can release energy. The key is that fission with energy release occurs only when very heavy nuclei break apart into medium-size nuclei, and fusion with energy release occurs when very light nuclei fuse into heavier ones. The answer again resides on the curve of binding energy; "mid-sized" elements have the highest binding energy per nucleon. As predicted by the curve, it requires energy to combine two medium-sized nuclei into a single large nucleus, just as it requires energy to split a medium-size nucleus into smaller nuclei.

Fusion is critical to the universe we observe. Without it, there would be no stars, and in fact no elements heavier than Lithium ($Z = 3$). However, there is also a practical reason for scientific interest in fusion. If we could create and control fusion, we could use it as an energy source. Using fusion as an energy source has two huge appeals: the fuel (hydrogen isotopes) is present in the oceans in essentially unlimited quantities (compared to the relative scarcity of, say, uranium) and fusion creates no radioactive byproducts.

However, before fusion reactors become commonplace, some daunting engineering challenges will have to be solved. Recreating the conditions inside the Sun, with its enormous temperatures, is no easy feat. Furthermore, simply getting the atoms hot is not sufficient; there is a minimum plasma density that must also be achieved so that collisions between nuclei will occur frequently enough to release energy, which keeps the plasma hot and keeps the fusion reaction going. The combination of high temperature and high density requirements necessarily means that the pressure must also be very high, to hold all of the reactants together.

As in the case of fission reactions, early work on fusion was directed toward nuclear weapons. Scientists solved the problem of achieving the ultrahigh temperature and pressure conditions necessary for fusion by using a fission ("atomic") bomb as a trigger, to both heat up and compress the fusion fuel. Some trigger! Using atomic bombs to power a nearby fusion reactor is not a very popular proposal.

There are two less-explosive schemes that are currently being pursued to keep the superhot plasma together for fusion to occur. *Magnetic confinement* uses electromagnetic fields to hold the charged particles together. *Inertial confinement fusion* uses a solid pellet of deuterium and tritium that is crushed by the light pressure of perfectly timed, short-duration, high-powered laser beams from all directions. The term comes from the fact that the particles' own inertia keeps them in place during the laser pulse.

Radioactive decay: A nucleus spontaneously emits particles or high-energy photons and either changes identity or becomes less excited.

Transmutation: Changing of one element to another after α or β radiation is emitted.

Fission is not the only way that a nucleus with an unstable combination of protons and neutrons can change into a more stable configuration. An unstable nucleus may instead spontaneously emit a charged particle or a high-energy photon in order to reach a more stable state. It changes via a *nuclear reaction*. Isotopes that change (decay) like this are said to be *radioactive*. It is possible for the isotope to become a different chemical element after the decay, a process known as *transmutation*.

The outgoing radiation can be classified as *alpha rays*, *beta rays*, or *gamma rays*. (They are represented by the Greek letters α, β, and γ.) Different decay processes result in different forms of radiation. Alpha and beta rays consist of matter particles, while gamma rays are photons (light particles).

In an alpha decay, the initial radioactive isotope decays into a different element by emitting an α particle. An α particle is made up of two protons and two neutrons. It is a helium-4 nucleus.

The initial isotope is known as the *parent*. Because it loses two protons, its atomic number is reduced by two (the mass number is reduced by four). Since the nucleus now has a different number of protons, it becomes another element − it has transmuted. The newly-formed nucleus is known as the *daughter*.

Alpha decay occurs most commonly in heavy nuclei whose ratio of protons to neutrons, Z/N, is too large, making them unstable. An α particle is a very stable particle, and the daughter nucleus that is left behind is more stable (tightly bound) than the parent. To put it another way, the net result of the radioactive decay is a reduction of the ratio Z/N.

Beta decay is characterized by the emission of an electron or antielectron. There are two types of beta decay, negative and positive. Negative β emission is represented in Concept 4. This occurs when a neutron inside the nucleus decays into a proton, an electron (the beta ray), and an almost zero-mass, uncharged particle known as an *antineutrino*. The antineutrino's interaction with matter is so weak that it is very hard to detect, and so it is customarily left out of nuclear equations.

The emitted electron did not exist in the nucleus beforehand, and is not one of the orbital electrons in the parent nucleus. When a neutron inside the nucleus turns into a proton, electron, and neutrino (and then emits the electron as a negative beta ray), its number of protons increases by one. The mass number stays the same. The released electron usually zips away, leaving behind a daughter atom with a net positive charge.

In positive β emission, the nucleus emits an antielectron (also called a positron). Antielectrons are essentially the same as electrons, except they have a change of **positive** e. The decay process is shown in Concept 5. A proton inside the atom decays into a neutron, an antielectron (the beta ray), and a neutrino. Like an antineutrino, a *neutrino* has no charge and almost zero mass. Its interaction with matter is also weak and it too is customarily left out of nuclear equations.

In emitting either type of β ray, the initial isotope changes atomic number, so β decay results in transmutation.

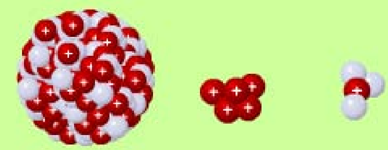
In *gamma decay*, the radioactive isotope emits a high-energy photon, also known as a γ ray.

Since the atomic number stays the same, the atom is the same chemical element after the decay. In this case it is the nuclear energy that changes in the decay. A nucleus can have different energy states. When a nucleus changes from an excited, high energy state to a lower one, a photon is emitted.

This is similar to the case when an electron falls from one energy level to another. A notable difference is that nuclear energy levels are much more widely spaced, on the order of millions of electron volts as opposed to say five or ten. This means the emitted photon, called a gamma ray, is much more energetic than the photon emitted by an excited atom.

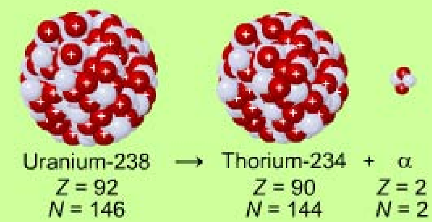
A gamma decay is represented in Concept 6. An excited nucleus is denoted by an

concept 1



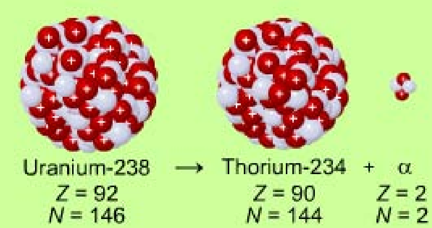
Unstable nuclei
Too massive or wrong ratio of protons to neutrons

concept 2



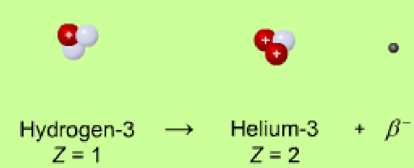
Radioactive decay with α or β ray
Nucleus decays by emitting charged particle
· Radioactive element has transmuted (changed to another element)

concept 3



Alpha radiation
Nucleus emits alpha particle (2 protons, 2 neutrons)
Number of protons in nucleus decreases by 2
· Transmutation
Mass number decreases by 4

concept 4

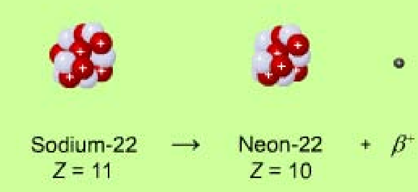


Negative beta radiation

asterisk “*” after the usual symbol. How does the nucleus get into an excited state? This usually happens after another kind of decay. In many cases of α and β decay, the product nucleus is in an excited state, after which it emits a γ ray and transitions to a lower state or to the ground state. Because the photon is electrically neutral, transmutation does not occur during gamma radiation.

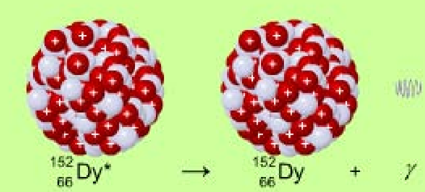
Nucleus emits negative beta particle (electron)
Number of protons in nucleus increases by 1
· Transmutation
Nuclear mass is unchanged

concept 5



Positive beta radiation
Nucleus emits positive beta particle (antielectron)
Number of protons in nucleus decreases by 1
· Transmutation
Nuclear mass is unchanged

concept 6



Radioactive decay with gamma ray
Nucleus emits gamma ray (high-energy photon)
· Becomes less excited
· No transmutation
· Nuclear mass is unchanged

43.16 - Representing radioactive processes

In this section, we look at the radioactive decay processes in more detail. Besides introducing the equations that are used to represent the processes, we emphasize the changes in proton number and mass number.

As seen in Concept 1, in an α decay, the parent turns into the daughter nucleus by emitting an α particle, which consists of two protons and two neutrons. This is a helium-4 nucleus, with $A = 4$ and $Z = 2$. As usual, the letter A represents the mass number (the total number of nucleons) and Z represents the proton number.

Here, P is used as a placeholder for the chemical symbol of the parent element, and the daughter element is represented by the D.

There are two types of beta rays. The equation for negative β emission is represented in Concept 2. This occurs when a neutron inside the atom decays into a proton, an electron (the beta ray), and an antineutrino (not shown in the equation). The electron is emitted, there is a newly-created proton in the nucleus, and the mass number stays the same.

The equation for positive β emission is shown in Concept 3. This occurs when a proton inside the atom decays into a neutron, an antielectron (the beta ray), and a neutrino (not shown in the equation).

concept 1

$$ {}^{A}_{Z}\mathrm{P} \rightarrow {}^{A-4}_{Z-2}\mathrm{D} + {}^{4}_{2}\alpha $$

Alpha radiation
Parent nucleus becomes daughter nucleus and an α particle
Number of protons in parent nucleus decreases by 2
Mass number of parent decreases by 4

In emitting α rays or either type of β ray, the initial isotope has a different proton number than the parent, so these decays result in transmutation.

In the final process we discuss here, γ decay, the radioactive isotope emits a high-energy photon, also known as a γ ray. This is shown in Concept 4. Since the atomic number of the nucleus stays the same, it is the same chemical element after the decay. It also has the same mass number since photons are massless. The only change is to the energy of the nucleus, which becomes less excited. Transmutation does not occur during gamma radiation, so the same symbol, X, is used to represent the element before and after the radiation occurs.

concept 2

$$^{A}_{Z}P \rightarrow {}^{A}_{Z+1}D + {}^{0}_{-1}\beta$$

Negative beta radiation
Parent nucleus becomes daughter nucleus and negative beta particle (electron)
Number of protons in parent nucleus increases by 1
Mass number of parent does not change

concept 3

$$^{A}_{Z}P \rightarrow {}^{A}_{Z-1}D + {}^{0}_{+1}\beta$$

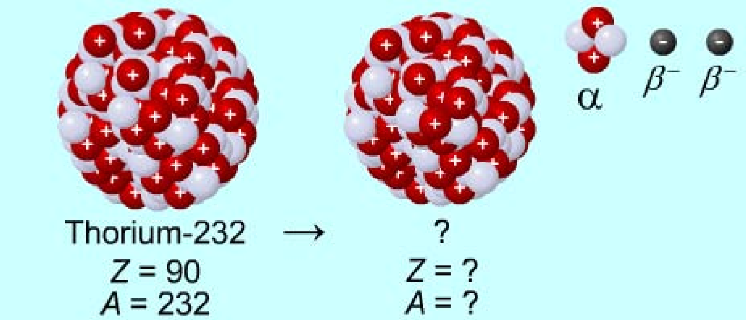
Positive beta radiation
Parent nucleus becomes daughter nucleus and positive beta particle (antielectron)
Number of protons in parent nucleus decreases by 1
Mass number of parent does not change

concept 4

$$^{A}_{Z}X^{*} \rightarrow {}^{A}_{Z}X + \gamma$$

Gamma radiation
Nucleus emits gamma ray (high-energy photon)
Number of protons in parent nucleus does not change
Mass number does not change

43.17 - Sample problem: radioactive decay



A thorium-232 nucleus (Z = 90, A = 232) is radioactive and decays by first emitting an alpha particle, then two negative beta particles. What are the atomic number and the mass number of the daughter nucleus after each of these three decay steps?

Variables

atomic number, number of protons	Z
mass number, total number of protons and neutrons	A

What is the strategy?

1. The α and β particles have known charge and known mass number. Subtract these from the atomic number and the mass number of the parent nucleus to determine the daughter nucleus's values of Z and A.

Physics principles and equations

An α particle consists of 2 protons and 2 neutrons. Its mass number is four.

A negative β particle consists of one electron. When the parent nucleus emits it, the number of protons increases by one and its mass number is unchanged.

Step-by-step solution

We will use the notation Z_0 and A_0 to denote the initial values of the proton number and the mass number, Z_1 and A_1 to denote their values after the first decay step, Z_2 and A_2 to denote their values after the second decay step, and Z_3 and A_3 to denote their values after the third and final decay step.

Step	Reason
1. $Z_0 = 90$ $A_0 = 232$	initial values
2. $Z_1 = Z_0 - 2$ $Z_1 = 90 - 2$ $Z_1 = 88$	α emission decreases nuclear charge
3. $A_1 = A_0 - 4$ $A_1 = 232 - 4$ $A_1 = 228$	α emission decreases nuclear mass

After the first decay process, the nucleus has 88 protons and 228 nucleons in total. This is radon-228, which then emits a negative β particle.

Step	Reason
4. $Z_1 = 88$ $A_1 = 228$	initial values
5. $Z_2 = Z_1 + 1$ $Z_2 = 88 + 1$ $Z_2 = 89$	β emission increases nuclear charge
6. $A_2 = A_1$ $A_2 = 228$	β emission does not affect nuclear mass

After the second decay process, there is a nucleus with 89 protons and 228 nucleons in total. This is actinium-228, which then decays by emitting another negative β particle.

Step	Reason
7. $Z_2 = 89$ $A_2 = 228$	initial values
8. $Z_3 = Z_2 + 1$ $Z_3 = 89 + 1$ $Z_3 = 90$	β emission increases nuclear charge
9. $A_3 = A_2$ $A_3 = 228$	β emission does not affect nuclear mass

After the third decay process, the nucleus has 90 protons, and a total of 228 nucleons. The end result after the three decays is again thorium because the number of its protons is again 90. After the three decays, the element is thorium-228. This lighter thorium isotope then continues to decay in a series of α and β decays until it becomes the stable isotope lead-208.

43.18 - Radioactive decay and half-lives

Half-life: The time it takes for half of a group of radioactive atoms to decay.

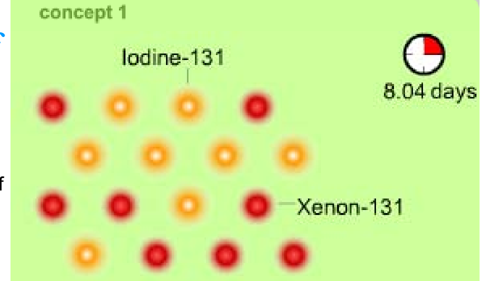
We have discussed the methods by which radioactive isotopes decay, by emitting α, β, or γ rays. What can be said about the rate at which they decay? If you have a sample of radioactive atoms, how many of them will decay in, say, the next minute?

The answer depends on the radioactive isotope. For any isotope, the fraction of atoms that will decay in a minute can be determined. A quantity of particular interest is the half-life, which is the average time it takes for one-half of the radioactive material present to decay.

For example, let's consider 16.00 mg of a parent nuclide, ^{131}I, which is a radioactive isotope of iodine. Its half-life is 8.04 days. You would find that after 8.04 days, one-half of the parent nuclei have decayed, and 8.00 mg of ^{131}I remains. After another 8.04 days, one-half of the remaining iodine will have decayed, and 4.00 mg remains. After a third 8.04 days, only 2.00 mg would remain, and after 8.04 more days, 1.00 mg would remain, and so on. Half of the remaining iodine decays every 8.04 days.

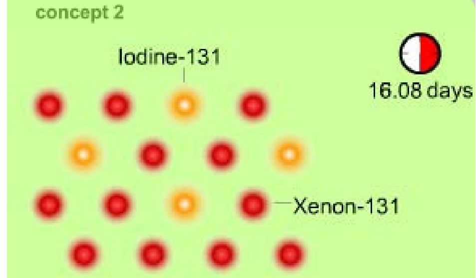
If it is possible to know how many atoms in a sample are going to decay within a certain time interval, is it possible to know when a particular atom will decay? Numerous experiments have shown that the answer to that question for a particular atom is no. This is similar to the situation of flipping one thousand coins and making a prediction of "50% heads." The prediction will be quite accurate, though you cannot reliably predict the outcome of any particular coin. The process of radioactive decay provided evidence of the statistical nature of quantum mechanics, which governs processes on a subatomic scale.

concept 1



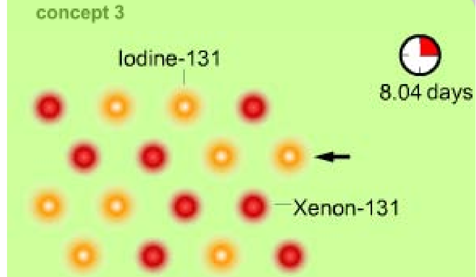
Half-life
Average time for one-half of a group of radioactive atoms to decay

concept 2



During each half-life, one-half of remaining radioactive atoms decay

concept 3



Radioactive decay is probabilistic
Cannot be predicted for an individual

atom
Can state the probability of any atom decaying within a certain time
· or what fraction of the atoms will decay within a certain time

43.19 - Radioactive decay equation

Numerous experiments have revealed the mathematical law that describes radioactive decay. If a certain fraction of the initial amount of radioactive material decays in a particular time interval, then the same fraction of the remaining atoms will decay in the next equal time interval, and the same fraction of the now-remaining atoms will decay in the next equal time interval, and so on.

Typically, this is discussed in terms of half-life, the time it takes half the material to decay, as shown in Equation 1. N_0 is the initial amount, N is the amount left at time t, and $\tau_{1/2}$ is a half-life, the average time it takes for half of the parent nuclei to decay.

Radioactive decay is steady over time, and does not depend on environmental variables such as pressure and temperature. This makes it possible to establish the ages of geological and biological samples by measuring the ratio of the amount of the parent nuclide to that of the daughter nuclide. The smaller this ratio, the older the specimen.

This technique is known as *radioactive dating*, and we illustrate how this works on biological samples with the method of *carbon-14 dating*.

Radioactive carbon-14 is created in the upper atmosphere by *cosmic rays*, which are showers of high-energy particles (mostly protons) that swarm throughout our galaxy. The carbon-14 radionuclide, ${}^{14}_{6}\mathrm{C}$, has too many neutrons to be stable and decays via β emission to ${}^{14}_{7}\mathrm{N}$, with a half-life of 5730 years. Although at any time, the supply of parent nuclide is diminishing due to decay, the cosmic-ray shower continually refreshes it.

Carbon-14 combines with oxygen to form carbon dioxide, which plants take in, and animals then eat. The mixing of carbon-14 with ordinary carbon-12 is very efficient, so living plants and animals all have the same ratio of ${}^{14}_{6}\mathrm{C}$ / ${}^{12}_{6}\mathrm{C}$. Once the organism dies, the plant/animal stops taking in new radioactive carbon atoms. However, the C-14 that has been ingested is still decaying into the daughter atom, nitrogen. After one half-life, for example, there is only one-half as much C-14 present (relative to C-12) as there was in living matter. By measuring this ratio, scientists can measure the ages of organic materials such as wood, cloth, leather, charcoal from cooking fires, and so on. This technique has proved to be a boon to fields such as archeology and biology.

equation 1

$$N(t) = N_0\left(\tfrac{1}{2}\right)^{t/\tau_{1/2}}$$

Radioactive decay equation

$N(t)$ = amount left at time t

N_0 = initial amount

$\tau_{1/2}$ = half-life of radioactive isotope

example 1

At 12 noon on Jan 1, this sample contains 1.25×10^12 atoms of iodine-131, whose half-life is 8.04 days. At 12 noon on Jan 31, how many atoms of iodine-131 remain?

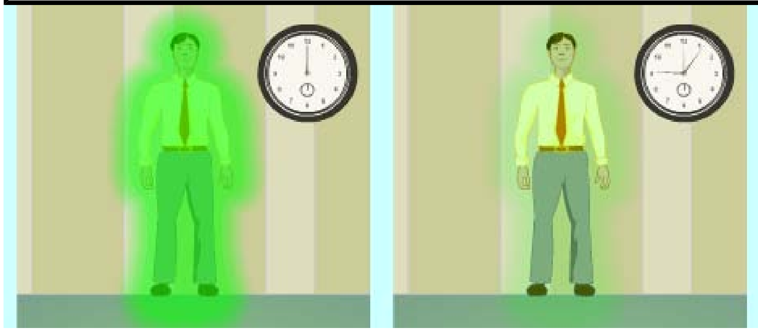
$$N(t) = N_0\left(\frac{1}{2}\right)^{t/\tau_{1/2}}$$

$$N(30\text{ d}) = (1.25\times 10^{12}\text{atoms})\left(\frac{1}{2}\right)^{\frac{30.0\text{ d}}{8.04\text{ d}}}$$

$$N(30\text{ d}) = (1.25\times 10^{12}\text{ atoms})\left(\frac{1}{2}\right)^{3.73}$$

$$N(30\text{ d}) = 9.42\times 10^{10}\text{ atoms}$$

43.20 - Sample problem: deducing the half-life



A patient is injected with radioactive technetium-99 and its progress is traced through the body. After 9.1 hours, 35% of the original isotope is left. What is the half-life of this isotope?

A technetium heart scan is a commonly performed, non-invasive nuclear scan that uses a radioactive isotope to evaluate blood flow after a heart attack. Metastable, excited technetium-99 nuclei decay via gamma emission, and a detector tracks the absorption of the isotope by the emitted radiation. In this problem, we assume that none of the radioisotope is lost through biological processes where it leaves the body; in other words, assume that the amount decreases only through radioactive decay.

Variables

initial number of radioactive technetium atoms	N_0
half-life	$\tau_{1/2}$
elapsed time	t
fraction of isotope left after t	35%

What is the strategy?

1. Apply the law of radioactive decay.
2. Substitute in the known information, and solve for the half-life.

Physics principles and equations

The law of radioactive decay

$$N(t) = N_0 \left(\frac{1}{2}\right)^{t/\tau_{1/2}}$$

Mathematics principles

A rule for the logarithm of an exponent.

$$\ln(a^b) = b\ln(a)$$

Step-by-step solution

Use the equation for radioactive decay, and solve for the half life.

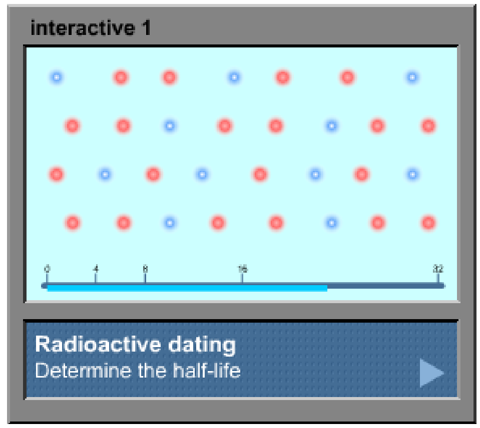
Step	Reason
1. $N(t) = N_0 \left(\frac{1}{2}\right)^{t/\tau_{1/2}}$	law of radioactive decay
2. $\frac{N(t)}{N_0} = \left(\frac{1}{2}\right)^{t/\tau_{1/2}}$	divide by the initial amount
3. $\ln\left(\frac{N(t)}{N_0}\right) = \frac{t}{\tau_{1/2}}\ln\left(\frac{1}{2}\right)$	take natural logarithm of both sides
4. $\tau_{1/2}\ln\left(\frac{N(t)}{N_0}\right) = t\ln\left(\frac{1}{2}\right)$ $\tau_{1/2} = \left(t\ln\left(\frac{1}{2}\right)/\ln\left(\frac{N(t)}{N_0}\right)\right)$	multiply both sides by $\tau_{1/2}$ and solve for the half-life

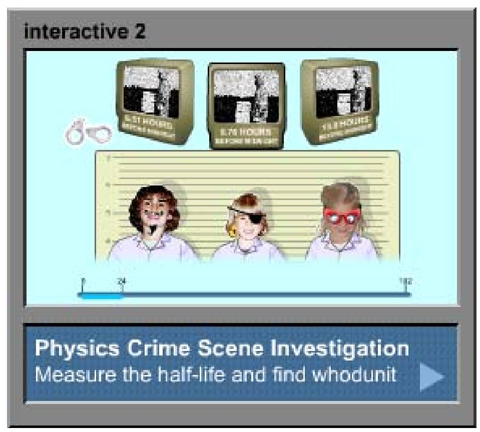
Now, use the given information and evaluate.

Step	Reason
5. $\tau_{1/2} = \left(t \ln\left(\frac{1}{2}\right) / \ln\left(\frac{N(t)}{N_0}\right)\right)$ $\tau_{1/2} = \left((9.1\text{ h}) \times \ln\left(\frac{1}{2}\right) / \ln(0.35)\right)$ $\tau_{1/2} = 6.0\text{ h}$	evaluate

The excited technetium-99 nucleus, which decays via gamma decay, has a fairly short half-life. This property makes it useful, because the radionuclide can be given in a large dose, which makes the detection process easier, while still keeping the patient's total exposure to radiation at a safe level.

43.21 - Interactive problem: radioactive dating





Carbon-14 is a radioactive isotope of carbon that has six protons and eight neutrons in its nucleus. It is commonly used to establish a date for organic specimens. In the first simulation, you will observe the decay of carbon-14 (C-14), and determine the half-life of that radioactive isotope.

You are equipped with a digital timer and a gauge that reports the number of parent atoms that are present. Before the simulation starts, there are 32 billion parent atoms. Each of the spheres on the screen represents a billion atoms. A sphere changes color when a billion atoms have decayed.

When you press GO, the timer starts and the carbon begins to decay to nitrogen by emitting β rays, and the number of daughter nitrogen atoms begins to grow. The daughter atoms are stable.

In the simulation, time is sped up and passes in thousands of years. The number of carbon atoms is shown both graphically with a "thermometer"-type gauge as well as numerically. When half the initial number of parent atoms has become daughter atoms, press PAUSE and note the elapsed time. (That is, at this moment, 16 billion of the carbon atoms remain present, and the rest have decayed.)

After noting the time, restart the process by pressing GO to resume the simulation and press PAUSE when half of the remaining parent atoms have become daughter atoms − in other words, when about eight billion carbon atoms remain. How does this second time interval compare to the first interval of time, when the number of parent atoms changed from 32 billion to 16 billion?

Repeat this process once more (or as many times as you like), each time pausing when the number of parent atoms has fallen in half. Do you see a pattern? You are measuring the half-life of the material, which is the average time needed for half of the parent atoms in a radioactive sample to decay. Enter your measurement for the half-life of carbon-14 by selecting the appropriate amount of years. Press CHECK to see if you are correct.

In the second simulation, you will use your newly acquired skills at measuring half-lives to investigate a crime scene. You are an environmental investigator and a criminal is once again dumping pure samples of a radioactive lead isotope, ${}^{209}_{82}\text{Pb}$, into a vacant lot. Holy ecological disaster!

You have been unable to catch the perpetrator in the act, but a security camera filmed three suspicious-looking characters in the vacant lot at different times. If you can determine when the radioactive waste was dumped in the lot, you will know which of these three suspects is guilty.

The factory that creates the waste is cooperating with you. They tell you that the isotope was pure lead-209 samples that initially contained 192 billion atoms. When you find the waste, it is about midnight. At midnight, 24 billion lead atoms remain, which means 168 billion of the lead atoms have decayed into bismuth atoms. In other words, one-eighth of the original lead-209 is left.

Your mission has three parts. First, determine how many half-lives have elapsed since the pure lead-209 sample was dumped. If you are having trouble with this piece of your detective work, return to the first simulation and calculate how many half-lives it takes for seven-eighths of the carbon atoms to decay.

Second, measure the half-life of lead-209 using a technique similar to what you used in the first simulation. You have the same tools as you had before.

Third, you can use the evidence from the security camera. The camera filmed Anna in the lot 6.51 hours before you obtained the sample. Sara was loitering in the area about 9.76 hours before this time, and a third suspect, Katherine, was filmed there 13.0 hours before the sample was found.

To put this all together: You use the value you determined for the half life of lead, and multiply that by the number of half-lives that have passed since the lead was dumped. That tells you how long the lead was there, so you can nail the suspect. To confirm your conclusion, drag the handcuffs in the simulation to the dastardly dumper. The simulation (and perhaps the suspect's reaction) will let you know if you are correct.

43.22 - Particle physics and GUTs

With this section, we come to the end of our discussion of the atom and the nucleus. It seems appropriate to take a peek ahead to the current state of nuclear physics, and to discuss what you and others may be learning in the decades ahead.

Particle physics, also known as high-energy physics, is one of the largest subfields of current physics. Essentially, it tries to answer the question "What is everything made of?" Ordinary matter is composed of the particles you have encountered so far − protons, neutrons, and electrons − and in the early 1900s, that seemed to be all that was needed to answer the big question. However, starting in the 1930s, particle physicists began discovering new, exotic particles that were created in an energy-to-mass conversion during collisions between the known particles. Several hundred other particles have since been discovered, most of which are unstable. Some of the particles are stable if they are left alone, but are composed of *antimatter*, which annihilates ordinary matter upon contact.

Early experimenters relied on cosmic rays (high-energy particles that permeate our galaxy) to initiate reactions. (The exact source of cosmic rays is still an open question. Stars emit them during intermittent flare-ups, but supernova explosions, when stars die, are thought to be responsible for much of the cosmic ray output.)

Later, more controlled experiments were carried out in particle accelerators, also known as atom-smashers, where particles are made to collide with higher and higher energies. Physicists now recognize that many of the heavier particles are made up from smaller building blocks called *quarks*. They understand that the smorgasbord of particles is due to the fact that when particles collide, newly-created quarks can combine with those already present to form systems of bound quarks.

The ultimate goal of physicists is a theory of everything (how is that for ambition?). Historically speaking, great breakthroughs in physics have often resulted in a simplification of our view of the universe. For example, Newton's universal law of gravity showed that the laws governing celestial orbits were the same as those governing the motion of objects falling under earth's gravity. Maxwell and his generation showed how electricity and magnetism, long thought to be unrelated, were really aspects of the same force. Physicists optimistically believe that the universe has an underlying simplicity, and that the number of fundamental forces can be reduced further.

We have discussed gravity, the electromagnetic force, and the strong force. (There is also the weak force which we have not discussed.) Albert Einstein spent a lot of his working life trying to interpret these forces as different aspects of a single "superforce". Historically, electricity and magnetism were united in the 1800s, and in the latter part of the 20th century, the weak force and the electromagnetic force were also joined theoretically. The goal of further reduction with the ultimate prize of unification still continues today.

As far as high-energy physicists are concerned, the goal is a *Grand Unification Theory*, or *GUT*. The current dream is to unify gravity with the strong, weak, and electromagnetic forces. You may have heard of attempts such as *superstring theory*, which interpret particles, such as electrons, as being modes of oscillation of unimaginably small "strings". Acceptance of superstring theory demands the idea that the universe may not consist of four dimensions (three spatial dimensions plus time), but instead ten or more dimensions. Bizarre as these ideas may seem, they are no less bizarre than relativity, nuclear theory, and quantum physics would have seemed to a scientist of the 19th century.

43.23 - Gotchas

All carbon atoms are the same. Not true: While carbon atoms all have six protons in their nucleus, different carbon atoms may have different numbers of neutrons. Nuclei with six protons but different numbers of neutrons are isotopes of carbon.

The nucleus is so small that an atom is about 99.99% empty space. No, but you are on the right track if you think this. Since the atomic diameter is about 10,000 times larger than the nucleus, the atom is more like 99.9999999999% empty space.

Protons and neutrons have different charge but approximately the same mass. Yes. A proton has a charge of 1.60×10^{-19} C and a neutron has no net electrical charge. The neutron is about 0.1% more massive than the proton, and they are each about 1800 times as massive as the electron.

The energy that is required to disassemble a nucleus into its constituent parts is the same amount of energy that is released when the same nucleus is assembled from separated nucleons. Yes; the **amount** of energy is the same and is called the binding energy. It requires energy to break apart a stable nucleus, and energy is released when the nucleus is assembled.

If half of a radioactive isotope decays during one half-life, then after two half-lives, it will all be gone. No, a half-life is the average time it takes for one-half of the radioactive material that is present to decay. After two half-lives, there will be one-quarter of the original amount left. After three half-lives, there will be one-eighth, and so on.

43.24 - Summary

Elements are substances that cannot be divided or changed into other substances using ordinary chemical methods. An atom is the smallest piece of an element that still has its chemical and physical properties. An atom consists of electrons orbiting a very small, very dense nucleus. The nucleus contains both protons and neutrons, which collectively are called nucleons.

The number of protons in an atom is the atomic number, Z. The number of neutrons is the neutron number, N. The total is known as the atomic mass number, A. Atomic masses are measured in terms of atomic mass units. The mass of a carbon-12 atom is defined to be exactly 12 u.

Isotopes are forms of an element with the same atomic number (which makes them the same element) but different numbers of neutrons (and hence different atomic mass numbers).

A fundamental, very short-range interaction called the strong force holds the nucleons together, counteracting the electrostatic repulsion between protons.

The nucleons are nearly incompressible and are tightly packed in the nucleus. The nuclear radius grows roughly as the cube root of the mass number.

Nuclei are stable only for certain combinations of protons and neutrons. On a plot of neutron number versus proton number, stable nuclei are represented as a band of stability that passes through the center of the diagram. For small nuclei, the numbers of protons and neutrons are roughly equal, but for large nuclei, the neutrons outnumber the protons by about 50%. Heavy nuclei need an excess of neutrons to dilute the proton concentration.

Binding energy is the energy that must be added to disassemble, or unbind, a nucleus into the protons and neutrons that make it up. The same amount of energy is released if the nucleus is assembled from nucleons that are initially separated.

The sum of the masses of the separate nucleons is always greater than the mass of the nucleus when it is whole. The mass difference is related to the binding energy by Einstein’s equation for mass-energy equivalence.

When the binding energy of one nucleus is greater than that of another, this means the particles are more tightly bound. For comparing how tightly bound two nuclei are, the binding energy per nucleon is the important figure.

The shape of the binding energy per nucleon vs. mass number curve is important. The curve is highest in the middle. This means that light nuclei can undergo the fusion process and release energy as this will increase their binding energy per nucleon. Similarly, heavy nuclei can undergo the fission process and release energy as this also will increase their binding energy per nucleon.

In the process of radioactive decay, an unstable nucleus spontaneously emits particles or high-energy photons. If the parent nucleus emits a charged particle such as an α particle, negative β, or positive β, then the number of protons in the nucleus changes. Transmutation is the changing of one element to another. If the parent nucleus emits gamma rays, which are uncharged photons, then transmutation does not take place, because the number of protons in the nucleus has not changed.

Equations

Z = atomic (proton) number

N = number of neutrons

Atomic mass number

$A = Z + N$

Atomic mass unit

$\mathrm{u} = 1.66054\times10^{-27}\ \mathrm{kg}$

Nuclear radius

$R = (1.2\times10^{-15}\ \mathrm{m})A^{1/3}$

Mass-energy equivalence

$E = mc^2$

Alpha radiation

${}^{A}_{Z}\mathrm{P} \rightarrow {}^{A-4}_{Z-2}\mathrm{D} + {}^{4}_{2}\alpha$

Negative beta radiation

${}^{A}_{Z}\mathrm{P} \rightarrow {}^{A}_{Z+1}\mathrm{D} + {}^{0}_{-1}\beta$

Positive beta radiation

${}^{A}_{Z}\mathrm{P} \rightarrow {}^{A}_{Z-1}\mathrm{D} + {}^{0}_{+1}\beta$

Gamma radiation

${}^{A}_{Z}\mathrm{X}^* \rightarrow {}^{A}_{Z}\mathrm{X} + \gamma$

Radioactive decay equation

$$N(t) = N_0\left(\frac{1}{2}\right)^{t/T_{1/2}}$$

Equations may be used to analyze the changes in atomic number and mass number after a radioactive decay.

The half-life of a radioactive isotope is the average time it takes for the decay of one-half of the atoms that are present in a sample.

The radioactive decay equation relates the amount of radioactive material that is present to the half-life, the elapsed time, and the original amount. This is the basis of the technique of radioactive dating.

Chapter 43 Problems

Conceptual Problems

C.1 In Rutherford's scattering experiment, positive alpha particles are directed at a thin gold foil. What was the key piece of evidence that led him to conclude that the positive charge and mass of the atom were concentrated in the nucleus?

i. All the alpha particles bounced straight back from the gold foil.
ii. None of the alpha particles were deflected by the gold foil.
iii. A small fraction were deflected from the gold foil.
iv. The alpha particles were absorbed by the gold foil.

C.2 What is the approximate ratio between an atomic diameter and a nuclear diameter?

i. 1:1
ii. 10:1
iii. 100:1
iv. 10000:1

C.3 Consider the following atoms: ^{12}C, ^{14}C, ^{14}N, ^{14}O, ^{16}O. The atomic numbers of carbon, nitrogen, and oxygen are 6, 7, and 8, respectively.
(a) Which atoms are isotopes of each other? Check all that apply.
(b) Which atoms have the same number of neutrons? Check all that apply.
(c) Which atoms have the same number of protons? Check all that apply.

(a) ☐ ^{12}C and ^{14}C
☐ ^{14}C and ^{14}N and ^{14}O
☐ ^{14}O and ^{16}O
☐ none of them

(b) ☐ ^{12}C and ^{14}C
☐ ^{12}C and ^{14}O
☐ ^{14}C and ^{16}O
☐ ^{12}C and ^{14}N
☐ none of them

(c) ☐ ^{12}C and ^{14}C
☐ ^{14}C and ^{14}N and ^{14}O
☐ ^{14}O and ^{16}O
☐ none of them

C.4 The strong force acts between which of the following types of particles? Check all that apply.

☐ proton and neutron
☐ proton and proton
☐ proton and electron
☐ neutron and neutron
☐ neutron and electron

C.5 What force or forces are responsible for keeping electrons in their orbits around the nucleus?

i. Strong force
ii. Electric (Coulomb) attraction
iii. Strong force and Coulomb attraction
iv. None of these

C.6 Consider the following atoms: ^{12}C, ^{14}C, ^{16}O. Rank these in order, from smallest to largest nuclear radii.

○ ^{12}C ^{14}C ^{16}O
○ ^{12}C ^{16}O ^{14}C
○ ^{14}C ^{12}C ^{16}O
○ ^{14}C ^{16}O ^{12}C
○ ^{16}O ^{12}C ^{14}C
○ ^{16}O ^{14}C ^{12}C

C.7 Can two iron atoms fuse to form a nucleus with greater binding energy per nucleon? Explain your answer.

○ Yes ○ No

C.8 Explain in your own words why a chain reaction is possible during uranium fission.

C.9 In the sun, 4 hydrogen nuclei undergo a multi-step process and eventually form a helium-4 nucleus (plus some positrons, neutrinos and gamma rays, all of negligible mass). What is this an example of?

i. Nuclear fission
ii. Nuclear fusion
iii. Alpha decay
iv. Chain reaction

C.10 A radioactive sample emits both alpha particles and negative beta particles. If a uniform external magnetic field is present in the region, what will be observed regarding the paths of the emitted particles?

i. Neither path bends.
ii. The paths of both particles bend in the same direction.
iii. Only one particle's path bends.
iv. The particles' paths bend in opposite directions.

Section Problems

Section 3 - Components of the nucleus

3.1 A neutral atom of potassium has 19 protons in its nucleus. How many electrons does it possess?

__________ electrons

3.2 For the neutral chlorine atom ${}^{35}_{17}\mathrm{Cl}$, determine the (a) mass number, (b) number of protons, (c) number of neutrons (d) number of electrons.

(a) __________ is the mass number
(b) __________ protons
(c) __________ neutrons
(d) __________ electrons

3.3 For a doubly-ionized calcium ion, which is a ${}^{40}_{20}\mathrm{Ca}$ atom with two electrons removed (leaving it with a net charge of $+2e$), determine the (a) mass number, (b) number of protons, (c) number of neutrons (d) number of electrons.

(a) __________ is the mass number
(b) __________ protons
(c) __________ neutrons
(d) __________ electrons

3.4 What is the mass of an electron, measured in atomic mass units?

__________ u

Section 6 - Nuclear properties

6.1 In a Rutherford-like experiment that analyzes electrons that are scattered from a sample, a scientist measures the radius of a nucleus in the sample to be 4.8×10^{-15} m. What of these is likely to be the mass number of the nucleus?

i. 1
ii. 2
iii. 4
iv. 16
v. 64

6.2 Find the mass number of a nucleus whose radius is approximately one-half that of uranium-240.

____

Section 7 - Sample problem: nuclear density

7.1 Estimate the radius of a sphere of nuclear matter whose mass is equal to that of the Earth, 5.97×10^{24} kg. Take 2.3×10^{17} kg/m^3 as the density of nuclear matter.

__________ m

Section 8 - Stable nuclei

8.1 Lead is an element that has 82 protons in its nucleus. Which of the following mass numbers is most likely to correspond to a stable isotope of lead?

i. 41
ii. 123
iii. 164
iv. 206
v. 246

Section 9 - Nuclear binding energy

9.1 What is the binding energy for a nucleus of helium-4, also known as an alpha particle? A neutron has a mass of 1.00866 u, a proton has a mass of 1.00728 u, and an alpha particle has a mass of 4.00153 u, where u = 1.66054×10^{-27} kg. Express the energy in joules to three significant figures.

__________ J

Section 13 - Fission

13.1 Consider the fission reaction that occurs when a neutron is captured by a uranium-235 atom and fissions into xenon-140, strontium-94, and 2 neutrons. Calculate the energy released in this reaction. The masses are: neutron = 1.0087 u, uranium-235 = 235.0439 u, xenon-140 = 139.9216 u, strontium-94 = 93.9154 u.

__________ J

Section 14 - Fusion

14.1 In the sun, 4 hydrogen nuclei undergo a multi-step process and eventually form a helium-4 nucleus (plus some positrons, neutrinos and gamma rays, all of negligible mass). For each helium atom produced, about 26 MeV of energy is released. (a) Which has more mass, the 4 hydrogen atoms or the helium-4 atom? (b) What is the magnitude of the mass difference?

(a) ○ 4 hydrogen atoms ○ 1 helium-4 atom
(b) __________ kg

14.2 The sun radiates electromagnetic energy at the rate of 3.91×10^{26} W. (a) What is the change in the sun's mass during one second? Express the answer as a positive mass. (b) The current mass of the sun is 1.99×10^{30} kg. What fraction of the sun's mass is lost in a century? Express the answer as a decimal number and assume years of 365.25 days.

(a) __________ kg
(b) __________

14.3 In the old days, people thought that the Sun might be a burning briquette of coal. At the Earth, the intensity of the radiation of the Sun is about 1400 W/m^2. Assume that the intensity is constant with time. The mass of the Sun is 2.0×10^{30} kg, and its distance from the Earth is 1.5×10^{11} m. The heat energy contained in coal is approximately 25 MJ/kg. How long would the Sun shine, if it really were a piece of burning coal?

__________ s

Section 15 - Radioactivity and radiation

15.1 An excited nucleus emits a gamma ray of energy 7.5 MeV. (a) Does the nuclear mass increase, or does it decrease? (b) What is the magnitude of the change?

(a) ○ Increase ○ Decrease
(b) __________ kg

15.2 A carbon-14 atom decays into nitrogen-14 (Z = 7). What particle(s) did it emit?

i. Positive beta particle(s)
ii. Negative beta
iii. Gamma
iv. Two alpha

15.3 A uranium-238 nucleus decays into a uranium-234 nucleus. What particle(s) did it emit?

i. Four positive beta particle(s)
ii. One alpha
iii. Two alpha and two negative beta
iv. One alpha and two negative beta

15.4 Lead (Z = 82, A = 214) decays in a series of steps. The nucleus emits 2 negative betas, 1 alpha, 1 negative beta, 1 alpha, and 1 negative beta. What is the final daughter product?

i. Lead-206
ii. Lead-205
iii. Lead-222
iv. Lead-210
v. None of these

15.5 Uranium (Z = 92, A = 234) decays in a series of steps. The nucleus emits 4 alphas, 1 negative beta, 2 alphas, 3 negative betas, and finally, 1 alpha. What is the final daughter product?

i. Lead-206
ii. Lead-208
iii. Lead-210
iv. Lead-212
v. Uranium-206

15.6 Uranium-238 (Z = 92) decays by alpha emission to thorium-234 (Z = 90), by emitting a helium-4 nucleus. Your friend claims to have witnessed a new type of decay where the uranium-238 emitted a helium-3 nucleus and became a thorium-235 atom. Prove that this is impossible for an isolated U-238 atom. Here is some helpful data: U-238 has an atomic mass of 238.051 u, the mass of a helium-3 nucleus is 3.014 u, and the mass of a Th-235 atom is 235.048 u.

Section 16 - Representing radioactive processes

16.1 Bismuth decays to form polonium in the decay represented by the equation

$${}^{212}_{83}\mathrm{Bi} \rightarrow {}^{212}_{84}\mathrm{Po} + \mathrm{X}$$

What particle is represented by X?

i. Proton
ii. Negative beta particle
iii. Alpha particle
iv. Gamma ray
v. Positive beta particle

16.2 Bismuth decays to form thallium in the decay represented by the equation

$${}^{213}_{83}\mathrm{Bi} \rightarrow {}^{209}_{81}\mathrm{Tl} + \mathrm{X}$$

What particle is represented by X?

i. Proton
ii. Negative beta particle
iii. Alpha particle
iv. Gamma ray
v. Positive beta particle

16.3 The radioactive element polonium (atomic number Z = 84) was used as a trigger in the original atomic bomb. The isotope of polonium with mass number A = 209 has a half-life of about 102 years, and decays by emitting an alpha particle. (a) Does the polonium transmute? (b) What is the atomic number of the daughter element? (c) What is the atomic mass number of the daughter element?

(a) ○ Yes ○ No
(b) __________
(c) __________

Section 18 - Radioactive decay and half-lives

18.1 Carbon-14 has a half-life of 5730 years. After four half-lives have elapsed, what percentage of an initially pure sample would remain unchanged?

__________ percent

Section 19 - Radioactive decay equation

19.1 The radioactive element cobalt-60 has a half-life of 5.27 years. An initially pure sample containing 7.32×10^{19} atoms is sealed in a time capsule. When the capsule is opened exactly 100 years later, estimate how many cobalt-60 atoms are left.

__________ cobalt-60 atoms

19.2 A scientist performing UN inspections finds evidence of an abandoned nuclear laboratory. A sample box in the lab reads "pure sample Strontium-90." However, the contents of the box are found to be only 61.0% Sr-90. The radioactive element Sr-90 is known to have a half-life of 28.6 years. How long ago was the sample created?

__________ years

19.3 The half-life of iodine-131 is 8.04 days. An eccentric scientist wishes to define a new quantity, the "two-thirds-life." For a radioactive sample, after a single two-thirds-life has elapsed, on average, one-third of the radioactive atoms will have decayed, and two-thirds will remain. What is the two-thirds-life for iodine-131?

__________ days

Section 21 - Interactive problem: radioactive dating

21.1 Using the simulation in the first interactive problem in this section, determine the half-life of Carbon-14.

i. 16000 years
ii. 11460
iii. 5730
iv. 2865

21.2 (a) Use the information for the second interactive problem in this section to determine the number of half-lives that elapsed between the time the pure lead-209 was dumped and midnight. (b) Use the second simulation to measure the half-life of lead-209. (c) Using the information given about when each suspect was spotted, determine who is the guilty party.

(a) __________ half-lives have elapsed
(b) __________ hours
(c) i. Anna
ii. Sara
iii. Katherine

Chapter 1 Answers

Answers to selected problems

C.3 No
Answers vary

C.5 Select all that apply.
$F = ma$
$v^2 = 2ax$
$F/v = m/t$

C.7 No
Answers vary

C.9 (a) Increases
(b) Decreases

2.1 100000 cm

2.3 9.966 teradollars

2.5 100000 tablets

3.1 3.75e−11

3.3 5.68e8 s

3.5 1.00e8

5.1 Dime

8.1 25.0 mi/h

8.3 1.28e6 cubits

8.5 Yes

8.7 (a) 5.79e10 m
(b) 5 910 000 000 km

8.9 2.72 m

8.11 4.9 kg

9.1 11.0 m/s

10.1 L/T^2

10.3 2

11.1 1.67e167

11.3 −1.98e5

12.1 (a) 1.01e10 dollars
(b) 3.75e3 dollars per person

12.3 1.6e−2 m/s^2

12.5 2e−12 %

13.1 5.59e7 kg

13.3 1.44e−3 kg

13.5 4.942e−8 mol

13.7 6.82e4 g

17.1 6.2 ft

17.3 8 cm

17.5 3.79 km

18.1 430 m

18.3 9.87 ft

18.5 330 m

19.1 1.88 radians

22.1 26.7 m/s^2

A.1 (a) 27.7°
(b) 0.483 rad
(c) 22.6 m

Answers to selected problems

C.1 No

C.5 (a) −15 m
(b) 15 m

C.7 Yes
Answers vary

C.9 Instantaneous velocity

C.11 Positive
Answers vary

C.15 Zero

C.17 No
Answers vary

C.19 Yes
Answers vary

0.1 Negative

1.1 1.6 meters

2.1 1 step(s)

2.3 (a) 5.3 km
(b) 5300 m

3.1 0.343 km

3.3 2.4e+3 s

4.1 (a) 0.063 m/s
(b) 240 m/s

4.3 (a) 138 km
(b) 110 km/h

4.5 (a) 10 ft/s
(b) 11 km/h

4.7 21.0 s

5.1 (a) 1440 m
(b) 7200 s

5.3 (a) 3.0 m/s
(b) −3.0 m/s
(c) 0 m/s

5.5 5.40 m/s

7.1 (a) Right
(b) Zero
(c) Start the ball moving slowly to the left and then increase its speed.

8.1 (a) 30.0 m/s
(b) −45.0 m/s
(c) 30.0 m/s

9.1 31 m

10.1 6.67 s

11.1 1.90e+5 m/s^2

11.3 (a) 0 m/s^2
(b) 9.6 m/s^2
(c) 7.7 m/s^2

11.5 (a) 0 cm/s^2
(b) −3.0 cm/s^2
(c) 3.0 cm/s^2

11.7 31v/5

12.1 (a) 1.0 s to 2.0 s
(b) 0 to 1.0 s
(c) 4.0 to 5.0 s

13.1 The hare

16.1 The white rabbit
The black rabbit

18.1 2.8 m/s

18.3 (a) 15.0 m
(b) 3.04 s

18.5 $a = 2(vt + d)/t^2$

18.7 773 m

22.1 −3.91 m/s^2

23.1 3.19 m/s

23.3 76.7 m/s

23.5 0.682 m/s^2

23.7 1.24 s

23.9 −4.2 m/s

23.11 (a) 1.22 s
(b) 9.10 m

23.13 25 percent

25.1 (a) 12 m/s
(b) −1.6 m/s^2

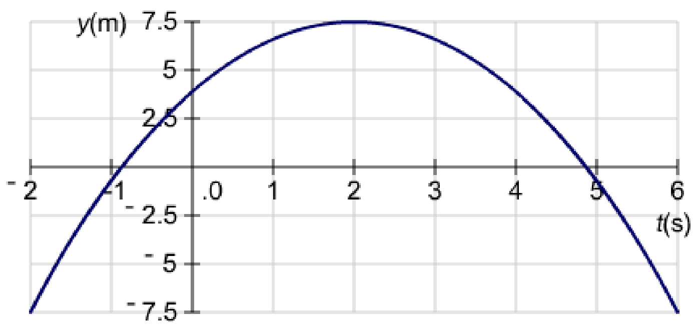
27.1 0.0481 s

29.1 1.70 m/s

A.1 (a) 3.5 m/s^2
(b) -1.8 m/s^2

A.3 (a) 4.01e14 m^3/s^2
(b) 9.49 m/s^2

A.5



(a) See above.
(b) 4.0 m
(c) 2.0 m/s
(d) 1.0 m/s
(e) 0 m/s
(f) −2.0 m/s^2

A.7 (a) 3.33 m/s
(b) 0 m/s^2
(c) 3.43 m/s
(d) 883 s

Answers to selected problems

C.3 6.0 knots Southwest

C.5 −**D**

C.7 Yes

0.1 (a) (−4.0 km, 3.0 km)
(b) (3.0 km, 2.0 km)
(c) (−3.0 km, −2.0 km)

1.1 81.2

1.3 1.41e3 kg/m^3

3.1 (7.00 knots, 121°)

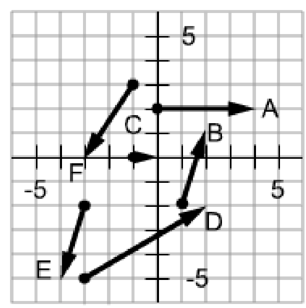
3.3 (3.00, 90.0°)

4.1 (a) (3, −25) blocks
(b) No

4.3 (a) **A** = (3, 2)
(b) **B** = (−1, −3)
(c) **C** = (−3, 3)

5.3 (a) Submit answer on paper.
(b) D
(c) C
(d) F
(e) B and E

(a)



6.1 (a) (18, 8)
(b) (3, 6)
(c) (10, 10, 10)
(d) (a+d,b+e,c+f)

6.3 (a) a = 5
(b) b = −4

9.1 (a) (18, −6, 48)
(b) (9, −12, 15)
(c) ($-a^2$, $-ab$, $-ac$)
(d) ($-2a$+18, −10−6b, $-2c$+12)

10.1 (a) (8, 230°)
(b) (21, 200°)
(c) (32, 80°)

11.1 (a) **A** = (2.72, 1.27)
(b) **B** = (−12.0, 12.0)
(c) **C** = (−3.98, 0.349)
(d) **D** = (3.98, 0.349)

11.3 (4.03, 4.97) cm

12.1 (a) **T** = (5.66 , 45.0°)
(b) **U** = (3.50, 90.0°)
(c) **V** = (5.83, 121°)
(d) **W** = (11.4, 232°)

12.3 (8.1 m, 210°)

16.1 (6000**i** + 6750**j** + 1500**k**) m

16.3 0**i** + −2**j** + −1**k**

16.5 (0.923, −0.231, −0.308)

17.1 (a) (5.0 km, 2.0 km)
(b) (3.6 km, −34°)
(c) −2.0 × (2.0 km, 1.0 km)

A.1 (32, 15, 5.0) km

A.3 (6.26 m, 21.9°)

Chapter 4 Answers

Answers to selected problems

C.3 On the Moon
Answers vary

C.5 (a) The speed of the river
(b) North
(c) No

0.1 (a) No
(b) The speed increases

1.1 (-8.5,−3.0,−1.0) m

1.3 (−17.0, −7.0) m

1.5 (4,5) m

2.1 (38.8,29.2) m/s

2.3 (108, 40.3) m/s

2.5 (a) 2.5733e+9 m
(b) Larger
(c) Smaller
(d) Answers vary

4.1 (a) (1.0, 4.0) m/s^2
(b) (0.0, 5.0) m/s

4.3 (180, -3.75) m/s

4.5 (0.73, 1.3) m

7.1 (a) 0.64 s
(b) 33 m/s

7.3 1.20 s

7.5 66.5 m

7.7 (a) 10.5 s
(b) 284 m
(c) −29.3 m/s
(d) 39.8 m/s

7.9 b

7.11 1.07 s

7.13 1 to 9

7.15 (a) 1 to 1
(b) Submit answer on paper.

9.1 At the professor's glove
Answers vary

11.1 -6.06 m/s

11.3 11.0 m/s

11.5 (2.36, 14.5) m/s

13.1 At the professor's glove
Answers vary

14.1 24.4°

14.3 (9.29, 20.6) m/s

14.5 35.3 m/s

14.7 (a) 1100 m/s
(b) 49.3°

14.9 4.07 s

14.11 (a) 1.93 m
(b) 2.29 s

14.13 (a) 2.03 m
(b) 218 m

18.1 30.2 m/s

19.1 3.1 m/s

20.1 (a) The smaller angle 26.3°
(b) The greater angle 63.7°

20.3 11.3 m

20.5 21.6°

22.1 8.0 m/s

23.1 40°

23.3 2.2 m/s

23.5 (a) Velocity of shortstop in runner's reference frame
(−15.0, 0) m/s
(b) Velocity of baseball in runner's reference frame
(−4.0, −18.0) m/s
(c) Velocity of runner in shortstop's reference frame
(15.0, 0) m/s
(d) Velocity of baseball in shortstop's reference frame
(11.0, −18.0) m/s
(e) Velocity of runner in baseball's reference frame
(4.0, 18.0) m/s
(f) Velocity of shortstop in baseball's reference frame
(−11.0, 18.0) m/s

Chapter 5 Answers

Answers to selected problems

C.1 Yes
Answers vary

C.5 Greater between blocks 1 and 2
Answers vary

C.13 Yes
Answers vary

0.1 It can be doing any of these

2.1 0 N (direction does not matter)

4.1 10 dumbells

4.3 (a) 209 N
(b) 21.3 kg

5.1 515 kg

5.3 (a) 5.0 s
(b) 7.0 s
(c) 6.0 s

5.5 1.2e−3 N

5.7 22500 kg

5.9 8.10e−7 N/m^2

5.11 (a) 0.0788 N
(b) 0.259 N

5.13 (a) 13.4 m/s^2
(b) 603 N
(c) Yes

5.15 5.1 m/s^2

9.1 (a) 6.54e4 N
(b) 7.57e4 N

10.1 −0.233 m/s^2

10.3 (a) Directly left
(b) 71.1 N

11.1 (a) 1.72 N
(b) 3.34 N

11.3 (a) 172 N
(b) 53.0°

12.1 117 N

15.1 (a) 49 N, Down
(b) 49 N, Up
(c) 78 N, Right
(d) 23 N, Left

17.1 The rightmost crate

19.1 2.03e3 N

19.3 260 N

20.1 0 m/s^2

20.3 0.13

20.5 0.0950

24.1 (a) 59 N, Straight down
(b) 51 N, Up and to the left
(c) 23 N, Down the plane
(d) 78 N, Up the plane

25.1 7.1 m/s

28.1 -0.301 m

28.3 (a)1.25 m/s^2
(b) Upward

28.5 (a) 4.12e+3 lbs/in
(b) 7.21e+5 N/m

30.1 1.60

31.1 17300 N

A.1 0.164

A.3 1.00e+3 N

A.5 (a) kg·s^2/m^3
(b) Submit answer on paper.

A.7 (a) Forward
(b) Backward
(c) Answers vary

Chapter 6 Answers

Chapter Assumptions

Unless stated otherwise, assume that all pulleys, ropes, strings, wires and other connecting materials are massless, frictionless and otherwise ideal.

Answers to selected problems

C.1 No
Answers vary

C.3 The bananas rise at the same rate and stay out of reach

C.5 Yes
Answers vary

0.1 (a) 12 N, Directly to the right
(b) Any magnitude greater than 10 N, Directly to the right
(c) 10 N, Directly to the right

1.1 127 N

1.3 20.9 N

1.5 (a) 26.3°
(b) 51.2°

3.1 (a) 1.06 m/s^2
(b) 330 N

5.1 180 N

6.1 648 N

7.1 685 N

7.3 0.323

7.5 (a) 83 N
(b) 0.14

7.7 17°

8.1 139 kg

9.1 4.90 m/s^2

9.3 (a) 8.54 m/s
(b) 12.0 m

11.1 (a) 0.0330 m/s^2
(b) 364 N

13.1 21200 N

A.1 $(F \cos \theta)/m$

A.3 (a) 6.79 **i** + −0.86 **j** N
(b) 2.09e−2 **i** + −2.65e−3 **j** m/s^2
(c) 9.41e−2 **i** + −1.19e−2 **j** m/s

A.5 (a) 0.305 m/s^2
(b) 1.90 N

Chapter 7 Answers

Answers to selected problems

C.1 No

C.3 No

C.5 No

0.1 (a) Greater
(b) To the same height as the other hill

1.1 201 J

1.3 181 J

2.1 -2.9e+3 m^2

2.3 37°

3.1 5.0 J

6.1 49 times higher

6.3 1410 J

7.1 (a) 8.0e−17 J
(b) 1.3e+7 m/s

7.3 1.63e+3 N

7.5 (a) 8460 J
(b) 26.0 m/s

10.1 133 N

12.1 2.1e+3 W

12.3 3.41 m/s

12.5 2.24e+5 J

13.1 $184

14.1 -8.58 J

14.3 2.04e+10 J

14.5 801 W

17.1 42.5 kg

17.3 1.44 m/s^2

19.1 63.8 m

20.1 0.400 m

22.1 8.0 m

24.1 2.7 m/s

25.1 9.4 m

27.1 No

27.3 (a) 8900 W
(b) 1200 N

A.1 (a) 131 m/s
(b) 6.3e+5 J

A.3 2.7 m/s

A.5 5.66 m/s

A.7 100 J

A.9 (a) 1.71e+10 J
(b) 146 s
(c) 7.80E+6 N

A.11 (a) −0.115 J
(b) 0.225 J
(c) −0.340 J
(d) 0.289

A.13 1820 W

Chapter 8 Answers

Answers to selected problems

C.1 No
Answers vary

C.3 (a) They have the same momentum
(b) A's kinetic energy is greater

C.5 Truck impulse equal

C.7 Center of mass
Answers vary

0.1 (a) Yes
(b) No
(c) Yes

1.1 1.29 kg·m/s

1.3 31 m/s

1.5 90 kg·m/s

2.1 2.8e2 kg·m/s

3.1 -1.73e+3 N

3.3 25800 N

3.6 (a) 1.13e3 kg·m/s
(b) 4.50e4 N

3.8 (a) 4.0 N
(b) 12 kg·m/s
(c) 80 m/s

3.10 2.79e+3 N

3.12 (a) −4.09e−23 kg·m/s
(b) 0.205 N

6.1 -1.26e+3 m/s

6.3 −3.25 m/s

6.5 (a) −0.43 m/s
(b) 0.77 m/s

10.1 32 N

10.2 118 kg

11.1 (a) 2.3 kg
(b) 3.1 m/s

11.3 (a) 8.1 kg
(b) −1.9 m/s

14.1 7.98e−3 m

15.1 6.0 m/s

16.1 (a) 38°
(b) 4.6 m/s
(c) No

16.3 (6.32 m/s, −22.3°)

16.5 (1.95 m/s, 26.8°)

17.1 1.39 m/s

18.1 (a) −0.40 m/s
(b) −3.6 m/s
(c) −1.8 m/s

18.3 0 m/s

18.5 2.0 kg·m/s

18.7 (a) 3.34 m/s
(b) No

19.1 11.8 m/s

19.3 0.012 m

20.1 4.7e6 m

20.3 (a) 6.0 m
(b) 6.3 m

22.1 831 meters

24.1 (a) Impossible
(b) Elastic
(c) Inelastic

A.1 (1.97, 11.4) m/s

A.3 (a) −1.20 m/s
(b) −15.1 J

Chapter 9 Answers

Answers to selected problems

C.7 Halfway as it moves from the lowest to the highest point

C.9 (a) The object's speed doubles
The radius of the circle is halved
(b) The object's speed doubles

C.3 (a) It is less on the Moon
(b) Equal

0.1 (a) Yes
(b) Yes

2.1 1.31e4 m/s

2.3 2.28e11 m

4.1 1.15 m/s^2

4.3 (a) 0.0337 m/s^2
(b) 0 m/s^2

4.5 21 m/s^2

5.1 9.0 m/s

6.1 4.50 "G's"

6.3 1.24e−4 N

6.5 6.3 N

7.1 7.8 m/s

7.3 21°

8.1 13.4°

8.3 48.4°

9.1 1.55 m/s^2

10.1 103 s

11.1 (a) 7.5 m/s
(b) 19 m/s

11.3 17.7 m/s

11.5 1.58 times the minimum velocity

13.1 (a) 8.0 m/s
(b) 9.5 m/s

A.1 8.61 m/s

A.3 10.5 m/s

A.5 4 to 1

A.7 (a) 0 m/s
(b) 34 m/s
(c) 17 m/s

A.9 3.07e3 m/s

A.11 11.1 m/s

Chapter 10 Answers

Answers to selected problems

C.1 Yes
Answers vary

C.3 Angular velocity
Angular acceleration
Answers vary

0.1 (a) Decrease
(b) Decrease
(c) Stay the same
(d) Increase

1.1 48 m

2.1 14 rad

2.3 7.17e−4 rad

2.5 (a) 3.54 s
(b) 0.315 rad
(c) 0.500 rad/s

3.1 2.0e+3 m

3.3 1.99e−7 rad/s

3.5 (a) 0.10 rad/s
(b) 23 rad/s

4.1 -28.0 rad/s^2

4.3 −0.12 rad/s^2

7.1 2.58 rev

7.3 0.0496 rad/s^2

7.5 4.6 m

7.7 1.33 s

7.9 (a) 8.33e−3 s
(b) 151 rad/s^2
(c) 300 rev

10.1 0.214 rad/s^2

11.1 (a) 7.08e−3 rad/s
(b) 2.01e−4 m/s^2
(c) 2.05e−5 *g*

11.3 17 m/s

11.5 0.619 rad/s^2

12.1 0.083 m/s^2

12.3 0.54 m/s^2

14.1 (a) 17 m/s^2
(b) 0°

14.3 (a) 9.78 m/s^2
(b) 4.55°

16.1 (a) 0.187 rad/s^2
(b) 6.00 m

A.1 (a) 190 rad/s
(b) 1.9e6 m/s

A.3 1.1 m/s

A.5 140 s

Answers to selected problems

C.2 (a) Yes
(b) Yes

C.4 False
Answers vary

C.7 To her left

C.9 (a) They walk at the same speed
(b) They move at the same speed

0.1 (a) Yes
(b) Farthest from the axis of rotation

1.1 376 N

1.3 0.21 N·m

2.1 15.0 N·m

2.3 2.75 N·m

3.1 14

3.3 (0, 0, −6)

4.1 1.60 rad/s^2

4.3 53 rad/s

5.1 19 kg·m^2

5.3 1.0e2 kg·m^2

6.1 1.3 s

6.3 2.0 rad/s^2

6.5 5.7e−5 kg·m^2

6.7 0.21

9.1 (a) −19 rad/s^2
(b) −4.6 m/s^2

10.1 2.85e5 N

11.1 (5/3)MR^2

11.3 0.15 kg·m^2

11.5 0.805 kg·m^2

12.1 325 J

12.3 835 J

13.1 1.25

13.3 4.95 rad/s

13.5 1.37 m/s

15.1 50.7 J

15.3 (a) 0.941 J
(b) 1.83 m/s
(c) 9.50 N

19.1 0.727 m/s^2

21.1 1.93e+5 kg·m^2/s

22.1 49.8 kg·m^2/s

22.3 0.47 kg·m^2/s

24.1 0.341 kg·m^2/s

26.1 143 rad/s

26.3 119 rad/s

27.1 50 rad/s

27.3 2.15 rad/s

29.1 0.120 m

A.1 2.0 kg·m^2

A.3 (a) 6.40 m/s
(b) −22.7 J

Chapter 12 Answers

Answers to selected problems

C.1 b
d

C.5 Steel,
Copper,
Titanium,
Aluminum

1.1 1.47 m

1.3 0.23 m

1.5 1.4e3 N

1.7 0.302 m

1.9 27 N

1.11 1.20 m

1.13 (a) 10.6 N
(b) 20.5 N

3.1 9.2e−2 m

3.3 (a) (1.0,0.50)
(b) (1.0,1.0)
(c) (0.63,0.38)
(d) (0.69,0)

3.5 0.827 m

3.7 0.32 m

4.1 (a) 145 N
(b) 349 N

7.1 (a) 950 N
(b) Negative *y*
(c) 1.32 m

11.1 (a) 1.6e5 N/m^2
(b) 2.0e6 N/m^2

12.1 0.0257 m

12.3 2.72e−5 m

12.5 4.0e9 N/m^2

12.7 82e9 N/m^2

13.1 1.5 m^3

13.3 −5.6e−5

13.5 −4.8e−7

14.1 4.90e+6 N/m^2

14.3 (a) 4.0e6 N/m^2
(b) 1.3e−4
(c) 7.9e−5 m

A.1 (a) 55.1 N
(b) 25.7 N

A.3 4.0e5 N/m^2

Chapter 13 Answers

Chapter Assumptions

Unless stated otherwise, use the following values.

r_{Earth}=6.38×10^{6}m

M_{Earth}=5.97×10^{24}kg

Answers to selected problems

0.1 (a) Decrease
(b) Increase

1.1 9.04e4 N

1.3 4.11e11 m

1.5 (2.0e−9,−5.2e−10) N

2.1 9.76 m/s^2

2.3 24.8 m/s^2

2.5 3.07e−5 m/s^2

9.1 7900 m/s

10.1 7.69e3 m/s

10.3 1.02e3 m/s

12.1 3160 m/s

13.1 (a) It increases
(b) It decreases

15.1 22.7 AU

15.3 (a) 1.00 AU
(b) 0.050

15.5 22.9 AU

18.1 (a) 3.74e8 s
(b) 11.9 years

18.3 2.40e+9 s

18.5 400 hours

18.7 1.05e4 years

20.1 (a) 42300 km
(b) 2610 m/s

21.1 (a) −6.32e11 J
(b) 3.16e11 J
(c) −3.16e11 J

21.3 1.47e+13 J

21.5 −1.14e10 J

25.1 9.1e12 J

26.1 1.04e4 m/s

26.3 (a) 3.55e4 m/s
(b) 3.76e4 m/s

26.5 (a) 1.47e4 m
(b) 2.95e12 m

A.1 (a) 3.58e7 m
(b) 3.07e3 m/s

A.3 49.9 solar masses

A.5 (a) 4.25e3 m/s
(b) 6.78e4 m/s

A.7 81.3 degrees

A.11 (a) 4.638e2 m/s
(b) 1.117e4 m/s
(c) 11 m/s

Chapter 14 Answers

Chapter Assumptions

Unless stated otherwise, use the following values:

Atmospheric pressure at the Earth's surface: $P_{atm}=1.013\times10^5$ Pa

Density of pure water = 1000kg/m^3

Density of seawater = 1030 kg/m^3

“Standard temperature and pressure” means 0°C and the atmospheric pressure stated above.

Answers to selected problems

C.3 No. The bottle will not explode.
Answers vary

C.5 No
Answers vary

C.9 The new waterline mark is lower on the hull.

C.11 The ball bearing
Answers vary

C.15 The lake's water level rises very slightly

2.1 (a) 31.3 people/km^2
(b) 42.7 people/km^2
(c) 1.86 people/km^2
(d) 6.10e3 people/km^2

2.3 (a) 12 kg
(b) 26 lb

3.1 (a) 6.85e6 Pa
(b) 1.57e5 Pa

3.3 (a) 7.64e4 Pa
(b) 1.95e3 Pa

3.5 52.5 N

4.1 (a) 1.01e7 Pa
(b) 5.05e7 Pa
(c) 1.01e8 Pa

4.3 (a) 10.3 m
(b) 0.760 m
(c) 760 mm
(d) 29.9 in

4.5 9.34e5 Pa

6.1 (a) 1.49e5 Pa
(b) 2.50e5 Pa
(c) Ben Chapman

7.1 (a) 1.01e3 kg/m^3
(b) 1.03e5 Pa

7.3 1.10e+5 Pa

9.1 6.76e+3 m^3

9.3 (a) 8.18e−2 m^3
(b) 1.05e3 kg/m^3
(c) 87.0 %

10.1 (a) 79.3 N
(b) 12.7 N

10.3 6.306e−2 N

12.1 159 N

12.3 (a) 4.50e4 N
(b) 4.19e3 m^3
(c) 5.26e4 N
(d) Yes

13.1 43.8 %

14.1 The 3.20 kg crown

15.1 63.8 N

18.1 36.6 m/s

20.1 8.36e+4 Pa

20.3 6.60e4 Pa

23.1 (a) 29.0 u
(b) 4.81e−26 kg
(c) 8.90e4 Pa at 1000 m
5.31e4 Pa at 5000 m
2.78e4 Pa at 10,000 m

24.1 (a) 29.9 Pa
(b) 8.57e−8 m^3/s

Chapter 15 Answers

Chapter Assumptions

The general form of the equation of motion for an object in SHM is $x(t) = A \cos(\omega t + \varphi)$.

Answers to selected problems

C.1 No
Answers vary

C.3 Remove mass

C.5 kg/s

0.1 (a) Stay the same
(b) Sinusoidal function

3.1 (a) 2.78e−4 Hz
(b) 2.31e−5 Hz

4.1 0.105 rad/s

5.1 (a) 5 m
(b) 3 m
(c) 4 m

5.3 $x(t) = 3.5 \cos((\pi/2)t + 2.0)$

6.1 (a) 0.40 m
(b) 2.0 s

7.1 $\pi/2 < \varphi < \pi$

7.3 2.2 rad

8.1 $x(t) = 7.8 \cos(5.6t + \pi)$

8.3 −1.9 m

9.1 103 m/s

11.1 0.15 m

13.1 (a) 0.011 m/s
(b) 0.071 m/s^2
(c) −0.051 m/s^2

13.3 3.85 s

14.1 $y(t) = 0.45 \sin((2\pi/1.2)t)$

15.1 0.73 s

15.3 (a) 0.753 m
(b) 3.43 s
(c) 0.858 s
(d) 1.38 m/s

15.5 3.12e+4 N/m

15.7 0.568 s

16.1 (a) 0.94 N/m
(b) 0.74 m/s^2

16.3 (a) 0.70 m
(b) 0.57 kg

18.1 13 N/m

19.1 0.030 J

19.3 3.7 J

20.1 2.1 m/s

21.1 670 J

21.3 0.836 s

22.1 0.018 $kg{\cdot}m^2$

23.1 3.15 m

23.3 90.0 minutes

24.1 1.20 m

25.1 1.0 s

26.1 1.3 s

26.3 (a) 0.967 s
(b) 1.01 s

Chapter 16 Answers

Answers to selected problems

C.1 Longitudinal
Answers vary

C.3 No
Answers vary

C.5 AM station
Answers vary

C.7 Wave on tauter string arrives first.

0.1 (a) Yes
(b) No
(c) No

4.1 2.5 m

5.1 6.0 m

6.1 0.00174 s

7.1 1.42e9 Hz

7.3 11 meters

7.5 0.497 m

8.1 347 m/s

8.3 1470 m/s

9.1 0.080 kg/m

10.1 331 N

11.1 −3.4e−2 m

11.3 (a) Right to left
(b) 3.34e−3 m
(c) 4.00 m
(d) 15.0 Hz

13.1 (a) 0.200 m
(b) 15.7 rad/m
(c) 6.28 rad/s

15.1 0.790 rad/s

A.1 (a) 0.20 Hz
(b) 0.60 m/s

A.3 0.20 s

A.5 (a) 189 N
(b) 2.90 m
(c) 156 m/s

Chapter 17 Answers

Chapter Assumptions

Unless stated otherwise, use 343 m/s for the speed of sound.

Answers to selected problems

C.5 The same

0.1 (a) Higher
(b) Stays the same

3.1 Higher

4.1 1.48e3 m/s

4.3 3.9 s

4.5 1.51e3 m/s

4.7 8920 kg/m^3

7.1 1.11e−5 m

8.1 12 m

8.3 1.72e−2 W/m^2

8.5 (a) 4
(b) 20

9.1 3.2e−5 W/m^2

9.3 50 dB

9.5 (a) 5.0e−17 W
(b) 1.6e−10 W
(c) 5.0e−3 W

9.7 20

12.1 787 Hz

12.3 44.6 m/s

12.5 33.1 m

14.1 745 Hz

14.3 (a) 4.93 kHz
(b) 5.19 kHz

14.5 (a) B
(b) 13.8 m/s

15.1 123 kHz

17.1 1.99 m/s

18.1 3.01

18.3 1.42

18.5 (a) 1.02e4 m
(b) 20.5 s

A.1 1.4 Hz

A.3 30 m/s

A.5 20 dB

A.7 260 m

Answers to selected problems

C.1 No
Answers vary

0.1 (a) Trough
(b) 2.00 m
(c) 1.00 m

3.1 (a) 3.3e2 m
(b) 2.0e3 m

3.3 (a) 20.9 rad/m
(b) 1150 rad/s

5.1 (a) 1.0 m
(b) 1.8 rad/m

7.1 97.1 Hz

7.3 2.29e+3 Hz

7.5 415 N

8.1 (a) 1.50 m
(b) 2

10.1 82 Hz

10.3 246 Hz

10.5 3.0e3

10.7 4.53 cm

13.1 (a) -2π
0
2π
(b) $-\pi$
π

14.1 1860 Hz

14.3 Destructive

17.1 5.36 Hz

17.3 4.4 Hz

17.5 1.28 Hz

17.7 0.665 m

Chapter 19 Answers

Answers to selected problems

C.1 (a) Answers vary
(b) Answers vary
(c) 99 °F

C.3 It does not change

2.1 (a) 0 °F
(b) 96 °F

3.1 (a) −173.15 °C
(b) −279.67 °F
(c) 22 °C
(d) −459.67 °F

3.3 119 °C

9.1 0.00196 m

9.3 2.062e−5 1/C°

9.5 158.3°C

10.1 4.48e8 N/m^2

12.1 0.103 cm^3

12.3 79.3 gallons

12.5 7.54e−3 m

12.7 −0.830 %

14.1 68.9 °C

14.3 2.57e7 J

14.5 7.6e−3 K

14.7 (a) 2.7 kg
(b) 52°C

14.9 1.68e6 J

15.1 538 J/kg·K

17.1 2.2 mol

19.1 1.37e+4 J

19.3 8.74e4 J

19.5 (a) 0°C
(b) 0.095 kg

23.1 (a) 2.7e2 W
(b) 47 W

23.3 (a) 8.59 in
(b) 23.7 in

23.5 2.4 W

24.1 0.27 W

24.3 21 °C

27.1 (a) 1.59e3 W
(b) 1.26e3 W

27.3 5.5e−5 m^2

27.5 586 W

29.1 2.70e5 J

Chapter 20 Answers

Chapter Assumptions

Use the following values for constants:

$N_A = 6.02\times10^{23}$

$R = 8.31$ J/mol·K

$k = 1.38\times10^{-23}$ J/K.

In problems which require you to know the atomic weights of atoms or molecules, use the following:

12.0 u for a carbon atom (C)

4.00 u for a helium atom (He)

1.00 u for a hydrogen atom (H)

14.0 for a nitrogen atom (N)

20.2 u for a neon atom (Ne)

16.0 u for an oxygen atom (O)

44.0 u for a carbon dioxide molecule (CO_2)

18.0 u for a water molecule (H_2O)

Answers to selected problems

3.1 1.73e5 Pa
3.3 0.150 m^3
3.5 4.80e−2 m^3

4.1 1.0e+22 molecules
4.3 (a) 44.0 u
(b) 4.40e−2 kg/mol
(c) 1.68e23 molecules

5.1 (a) 7.2 moles
(b) 4.3e24 molecules
5.3 1.24e5 Pa
5.5 −54.5 mol
5.7 741 K

7.1 2.00e−22 m^3

10.1 1.1e−20 J
10.3 531 K

11.1 31.0

12.1 377 m/s
12.3 450 m/s

Answers to selected problems

C.1 Adiabatic

0.1 (a) The heat is equal to the change in internal energy.
(b) The work done is the negative of the change in internal energy.
(c) The net heat transferred is equal to the sum of the change in internal energy and the work done.

1.1 (a) −3.4e5 J
(b) −5.1e5 J
(c) 1.7e5 J

1.3 −1700 J

3.1 (a) 9.50e3 J
(b) Work is done by the system

3.3 (a) 5.48e5 J
(b) 7.15e5 J

7.1 (a) 986 J
(b) Work is done by the system

10.1 −245 J

11.1 20 K

11.3 4.6 mol

15.1 (a) −3.2e3 J
(b) Work is done on the system

15.3 1.2e5 Pa

16.1 5.8e+3 J

16.3 231 mol

18.1 3.42 mol

19.1 350 K

19.3 4.16e+5 Pa

19.5 3.17

20.1 1.32 m^3

20.3 (a) 695 J
(b) Work is done by the system

23.1 58 J

24.1 (a) 7500 J
(b) −4500 J
(c) −2080 J

A.1 (a) 593 K
(b) 5.25e3 J

A.3 3.94e3 J

A.5 (a) 5.71e3 J
(b) 3.96e3 J
(c) 1.75e3 J

Chapter 22 Answers

Answers to selected problems

C.1 Stays the same

C.3 The room temperature goes up
Answers vary

C.4 Isothermal

C.6 The entropy of the system is unchanged.

1.1 177 J

1.3 930 J

5.1 0.222 J/K

8.1 1.45 J/K

8.3 674 J

10.1 458°C

12.1 (a) 755 K
(b) 1.21e+5 J

12.3 8.88e+3 J

12.5 (a) 28.7 %
(b) 53.2 %

13.1 (a) 47.6 %

13.3 (a) 1.22
(b) 1.33

15.1 (a) 16
(b) 88 J

15.3 (a) 1.4e5 J
(b) 97 cents

17.1 0.0281 L

A.1 (a) 1.52 J/K
(b) 0 J/K

A.3 (a) 0.969 J/K
(b) 0 J

Answers to selected problems

C.1 Richard
Answers vary

C.3 (a) 0 C
(b) 0 C

C.5 (a) Rubber is a(n) insulator.
(b) Iron is a(n) conductor.
(c) Copper is a(n) conductor.
(d) Wood is a(n) insulator.

C.7 the "10 cm" end

C.9 The force does not change.

C.11 (a) Yes
(b) Answers vary

C.13 (a) No
(b) Yes

0.1 (a) Away from each other
(b) Towards each other
(c) Decreases

1.1 +3.20e−19 C

1.3 1.60e−20 kg

2.1 (a) 0 C
(b) −1.60e−19 C

2.3 3.86e+13 electrons

2.5 1.5e20 electrons

3.1 (a) $-e$
(b) π

5.1 0 C

7.1 −3.6e−6 C

8.1 (a) Answers vary
(b) +4.00 μC
(c) 40000 dollars

9.1 2.0e−5 C

9.3 (a) 2.2 N
(b) 5.5e−5 C
(c) 45 kg

10.1 (a) Attractive
(b) 8.22e−8 N

10.3 9.32e−7 C

12.1 0.072 N

12.3 3.02e3 N

12.5 (a) NW corner −Q
SW corner +Q
SE corner −Q
(b) NE corner −Q
NW corner −Q
SW corner +Q
SE corner +Q
(c) NE corner +Q
NW corner −Q
SW corner −Q
SE corner +Q

12.7 −167 C

12.9 3.72e−2 N, directed 42.6° below the x−axis.

12.11 The unit vector has components 0.845 in the up, 0.169 in the east, and 0.507 in the north directions.

16.1 34 protons

A.1 (a) 0 C
(b) 2.27e39
(c) 3.79e12 kg
(d) 3.36e−39 C

A.3 (a) −1.92e−9 C^2
(b) −3.10e−5 C
(c) 6.20e−5 C
(d) 1.70 m/s
(e) No
Answers vary

Chapter 24 Answers

Answers to selected problems

C.1 2

C.3 (a) The combined field is zero
(b) To the right
(c) The combined field is zero

C.5 Between the protons
Right of the electron

C.7 Proton
Answers vary

0.1 Greatly increases

1.1 (a) 1.75e−3 N
(b) −2.89 N
(c) −184000 N
(d) 1.87e−12 N

1.3 1.36e−21 N

1.5 600 N

2.1 2.61 N/C

2.3 (a) 5.14e11 N/C
(b) Away from the proton

2.5 (a) 3.53e6 N/C
(b) 5.65e7 N/C
(c) 2.82e7 N/C

5.1 (a) Closer together
(b) The direction of the force is tangent to the field lines

6.1 (a) 2.59e11 N/C, to the right
(b) 1.12e11 N/C, to the left
(c) 9.55e10 N/C, to the right

6.3 16.2 C

7.1 4.14e−5 m

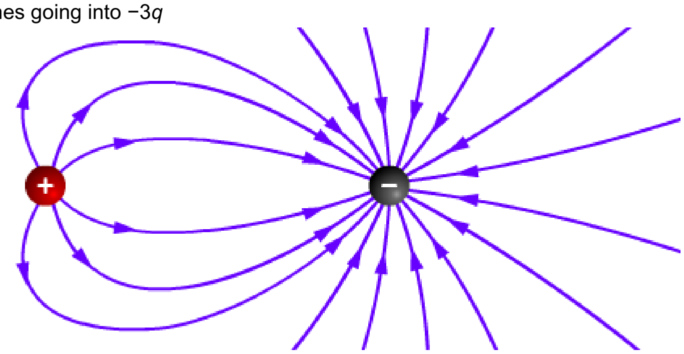
7.3 5.46e5 N/C, 250 °

8.1 7.17e+11q N/C, 225°

8.3 −1.62e10 N/C, 270°

8.5 (1.01e10, 1.97e10) N/C

9.1 18 lines going into −3q



9.3 (a) q_2
(b) q_1 and q_3
(c) $|q_2| > |q_1| > |q_3|$
(d) 3 times stronger

10.1 89.9 s

10.3 (a) −8.69e−17 N
(b) −9.54e13 m/s^2
(c) 1.00e6 m/s
(d) 3.33e−3

10.5 4.64e3 m/s

10.7 (a) 4.79e10 m/s^2
(b) 7.31e−7 s
(c) 1.28e−2 m
(d) 1.02e−18 J

12.1 1760 times larger

12.3 (a) 5.81e−4 s
(b) 8.72e−3 m

14.1 4.94e+4 m/s

15.1 2340 N/C

17.1 9.76e−9 C·m

17.3 6.59 m

A.1 (a) 4.7e+6 N/C
(b) 8.3e−3 C

A.3 (a) Negative
(b) 4.55e−3 N Away from the Earth

A.5 (a) 4.00e−13 N
(b) 8.00e−13 N

A.6 9.00e−6 N/C

Chapter 25 Answers

Chapter Assumptions

Unless stated otherwise, the reference configuration for zero electric potential and zero electric potential energy is one in which there is infinite separation between charges.

Answers to selected problems

C.1 (a) It doubles
(b) Yes
(c) No

C.5 (a) No
Answers vary
(b) Yes
Answers vary

C.7 a

C.9 (a) No
(b) Answers vary
(c) No
Answers vary

0.1 (a) Always positive
(b) Closer to

0.3 (a) It increases
(b) A vertical line

1.1 (a) 5.4 μJ
(b) −5.4 μJ
(c) Closer together

1.3 (a) −52.3 J
(b) 52.3 J
(c) Farther apart

2.1 (a) 5.09e−3 J
(b) −5.09e−3 J
(c) 5.09e−3 J

2.3 (a) 2.75e−2 J
(b) −2.75e−2 J
(c) 2.75e−2 J

2.5 (a) 6.71e−4 J
(b) −6.71e−4 J
(c) 6.71e−4 J
(d) Yes

3.1 2.6 J

3.3 (a) 2e2 N
(b) 2e−13 J

3.5 3.0e−12 m

3.7 4.48e6 J

4.1 (a) 2
(b) 1

5.1 (a) 4.89e7 J
(b) −7.46e7 J

5.3 (a) −2.46e−2 J
(b) It stays the same

7.1 (a) 1.44e−9 V
(b) −1.44e−9 V
(c) Near the proton
(d) Magnitudes are equal

7.3 2.41e−9 C

7.5 −3.26e−4 J

7.7 1.31e6 m/s

8.1 (a) −1.57e+6 J
(b) 3.60e+8 V

8.3 (a) 8.99e8 V
(b) 8.99e8 V

8.5 8.27e5 V

8.7 2.8e11*q*

11.1 32 V

12.1 (a) 6.34e−20 J
(b) −6.34e−20 J
(c) 0.396 V

12.3 (a) 1.28e−20 J
(b) 8.83e25 ions

12.5 (a) 1.04e4 J
(b) $ 2.6e−4

12.7 −2.66e−2 C

14.1 (a) −1.28e3 V
(b) 61000 m/s

14.3 (a) 2.00e−14 J
(b) 2.00e−14 J

16.1 2.53e+8 V

16.3 (a) Sphere
(b) The surface closer to the point charge
(c) 0.32

17.1 −9.6e−14 J

17.3 -40.9 V

17.5 710 V

17.7 (a) Planes parallel to the *yz* plane
(b) x = 2.24 m
(c) x = 1.87 m

Answers to selected problems

C.1 The flux through the spine

C.3 A capped bottle of root beer with infinitely thin walls

C.5 A sphere

C.7 The flux stays the same
Answers vary

0.1 (a) Stays the same
(b) Increases

1.1 (a) They have equal magnitude but opposite signs
(b) 0 (N/C)·m^2

1.3 (a) 14.0 N/C
(b) 0.180 N/C

1.5 4.5 (N/C)·m^2

1.7 (a) 0 m^2
(b) 0 m^2

1.9 (a) 1.84e−2 m^3/s
(b) 1.30e−2 m^3/s

3.1 (a) The flux doubles.
(b) Flux is directly proportional to surface area.
(c) The flux doubles.
(d) Flux is directly proportional to electric field strength.

3.3 2.7 m^2

5.1 (a) −1.1e5 (N/C)·m^2
(b) 0 (N/C)·m^2
(c) 1.1e5 (N/C)·m^2
(d) −1.1e5 (N/C)·m^2

5.3 (a) 6.02e−7 C
(b) There is a single point charge
The charge is spherically symmetric

5.5 (a) 1.43e−2 (N/C)·m^2
(b) −1.43e−2 (N/C)·m^2
(c) 1.43e−2 (N/C)·m^2
(d) 0 (N/C)·m^2

7.1 (a) 2e−9 C
(b) 5e−9 C

7.3 +2 nC; +3 nC

8.1 3.89e-11 C/m^2

8.3 2.95e6 m/s

8.5 (a) 9300 protons
(b) 3.34e−12 N/C
(c) 49.6 m^2
(d) 37200 protons
(e) 3.34e−12 N/C

8.7 (a) In the positive x direction.
(b) In the negative x direction.
(c) $\rho x/\varepsilon_0$
(d) $\rho h/2\varepsilon_0$

9.1 (a) 1.88 N/C
(b) 7.50e−2 N/C
(c) 0 N/C

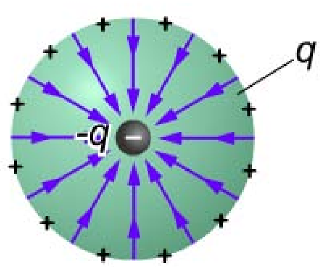
9.3 4.19e−10 C

9.5 (a) 0 N/C
(b) $(3.00\times10^{-3})k/r^2$
(c) See below.

10.1 (a) 5.69e6 N/C
(b) 5.69e6 N/C
(c) Answers vary

A.1 1.1e−10 C

9.5 Answer to (c)



Chapter 27 Answers

Chapter Assumptions

For any problems in this chapter that involve a household electrical outlet, assume that the voltage supplied is fixed at 120 V. While this is not true for actual outlets, whose voltage varies and periodically exceeds 120 V, the simplification will not affect the accuracy of your answer.

Answers to selected problems

C.1 (a) 1
(b) 3

C.3 The current stays the same
Answers vary

C.7 Non−ohmic

C.9 No
Answers vary

C.11 No
Answers vary

C.13 Electric energy

1.1 (a) 6.00e−2 C
(b) 3.75e17 electrons
(c) 0 protons

1.3 12.3 C

1.5 7.4e-7 A

1.7 4.2e−7 s

1.9 (a) 3.89e−2 C/m^2
(b) Down

1.11 54.6 A

2.1 (a) 1.1e4 s
(b) Answers vary

3.1 16.4 Ω

3.3 4.0e−3 A

3.5 1.29 V

4.1 100 Ω

5.1 2.4e−7 Ω

5.3 Aluminum

5.5 (a) 10.0 Ω
(b) 16.0 Ω
(c) 2.00 Ω

5.7 (a) 7.14e−7 m^3
(b) 5.75e−8 m^2
(c) 12.4 m

7.1 360°C

7.3 (a) 0.00501°C^{-1}
(b) Tungsten

7.5 16.3°C

7.7 51.8 Ω

8.1 3.1 V

8.3 20 light bulbs

8.5 14 Ω

8.7 (a) 2.4 Ω
(b) 1.0e2 W

9.1 $7.17e20

13.1 2.07e10 W

Answers to selected problems

C.1 (a) It increases
(b) It stays the same

C.2 (a) Increased
(b) Same amount of charge

0.1 Charge is proportional to potential difference

1.1 3.2e−10 C

1.3 7.1e−6 F

1.5 7.2e−5 C

3.1 1.5e−9 F

3.3 3.22e−3 m

5.1 7.74e−3 m

7.1 4.00e−9 F

7.3 4.45e−1 J

7.5 9.96e−7 F

7.7 1.54e−2 J

9.1 1.01e8 V

10.1 6.9 J/m^3

10.3 0.774 m

10.5 7.11e9 J/m^3

11.1 76 J

13.1 7.2 nF

13.3 1.60e8 V/m

13.4 66.0 nF

16.1 1.6e−7 C

17.1 3.1e−7 F

A.1 5.94e−10 F

A.3 8.15

Chapter 29 Answers

Answers to selected problems

C.5 The same
Answers vary

C.7 Circuit with two light bulbs
Answers vary

C.9 Parallel
Answers vary

C.13 (a) R_1 is in series with the combination of R_2 and R_3
(b) R_2 is in parallel with R_3

C.15 It will decrease to zero
Answers vary

C.17 (a) 2.0 V
(b) 3.0 V
(c) 4.5 V
(d) 4.5 V
(e) 3.0 V

C.19 (a) No
(b) No
(c) Answers vary

0.1 (a) They are equal
(b) It is the same everywhere

2.1 1.17e6 J

3.1 3.0 V

3.3 (a) 6.6 W
(b) 3.6e4 J

3.5 2.0 Ω

3.7 1.8e4 s

4.1 6.5 Ω

4.3 8.99 V

4.5 0.625 A

4.7 (a) 4.27 Ω
(b) 2.39e−2 Ω

6.1 0.55 A

7.1 1.0 Ω

7.3 4200 Ω

7.5 3.23 Ω

7.7 2.7 A

7.9 12.0 V

8.1 25.0 Ω

11.1 1.67 Ω

11.3 0.175 Ω

11.5 77 Ω

11.7 8.75 Ω

11.9 34.8 V

11.11 4

11.13 (a) 500 Ω
(b) 125 Ω

12.1 60.0 Ω

13.1 (a) 12.0 V
(b) 0.300 A
(c) 20.0 Ω

14.1 (a) 3.75 Ω
(b) 5.00 Ω

14.3 4.17 Ω

17.1 2.8 V

17.3 0.105 A

19.1 (a) 30.0 Ω
(b) 8.33e−3 A

20.1 (a) 0.6 A
(b) 12 Ω

20.3 5.81 A

23.1 (a) 20.0 Ω
(b) 20.0 V

24.1 (a) 5.00 ohm bulb
(b) Zero−resistance wire
(c) Zero−resistance wire
(d) 5.00 ohm bulb
(e) 20.0 ohm bulb

26.1 3.3e−5 F

26.3 23.4 V

26.5 (a) 4.55e−6 F
(b) 5.46e−5 C
(c) 4.20 V

27.1 1.5e−4 F

28.1 70 μF

28.3 (a) 20.0 μF
(b) 8.40e−5 C
(c) 12.0 V

28.5 $(11/5)C_1$

28.7 11

28.9 6.00e−6 F

31.1 869 Ω

31.3 3.03e4 Ω

31.5 5.19e−9 F

31.7 (a) 7.20e−5 C
(b) 1.44e−4 C
(c) 2.16e−4 C

32.1 24 s

32.3 (a) 1140 A
(b) 213 A
(c) 3.18e−6 C

32.5 (a) 2.4 A
(b) 7.2 V

Chapter 30 Answers

Chapter Assumptions

Elementary charge, $e = 1.60\times10^{-19}$ C

Mass of electron, $m_e = 9.11\times10^{-31}$ kg

Mass of proton, $m_p = 1.67\times10^{-27}$ kg

Unless stated otherwise, use 5.00×10^{-5}T for the strength of the Earth's magnetic field at its surface.

Answers to selected problems

C.1 Down
Answers vary

C.3 (a) No
(b) Electron

C.5 No
Answers vary

C.7 (a) Toward you
(b) Left
(c) Down
(d) Down

C.9 (a) Negative
(b) Positive
(c) Answers vary

C.11 Away from you
Up

C.17 (a) 180°
(b) 0°
(c) 90°
(d) 180°

0.1 (a) Curve
(b) No
(c) Yes

7.1 (a) 6.08e−16 N
(b) Positive *y* direction

7.3 (a) 805 m/s
(b) East

7.5 1.78 C

7.7 17.3°

7.9 3.60e6 m/s

9.1 (a) 2
(b) 180°
(c) 90.0°

10.1 0.417 T

10.3 0.631 T

10.5 5.76e−3 T

11.1 2.90 T

12.1 3.00e+4 m/s

12.3 2.21e−22 J

12.5 (a) 1.0e4 m/s
(b) 1.8e−2 m/s

13.1 2.75e−3 m

13.3 56.3 m/s

13.5 3.67e+2 s

13.7 2.28e−12 T

16.1 (a) 1.29e−2 T
(b) Decrease

16.3 (a) 4.65e−3 m
(b) 5.65e−3 m
(c) 2.00e−3 m

16.5 0.0286 m

18.1 Xenon

19.1 (a) 8.29 m/s
(b) 1.07 m
(c) 0.872 m/s
(d) 0.708 m

21.1 (a) A circle
(b) A straight line

22.1 14.1 N

22.3 2.27 A

22.5 35.2°

24.1 1.41e−5 N·m

24.3 2.66e−2 kg

25.1 0.25 A·m^2

25.3 2.89e−2 m

25.5 (a) τ = 3.68e−3 N·m, PE_B = −2.13e−3 J
(b) τ = 2.73e−3 N·m, PE_B = 3.26e−3 J

25.7 7.78e−3 A·m^2

A.1 2.17 rad/s

Chapter 31 Answers

Chapter Assumptions

Elementary charge, $e = 1.60\times10^{-19}\ \mathrm{C}$

Mass of electron, $m_e = 9.11\times10^{-31}\ \mathrm{kg}$

Mass of proton, $m_p = 1.67\times10^{-27}\ \mathrm{kg}$

Unless stated otherwise, use 5.00×10^{-5}T for the strength of the Earth's magnetic field at its surface.

Answers to selected problems

C.1 (a) The field is perpendicular to the plane of the wires.
(b) Reinforce
(c) Cancel
Answers vary

C.3 No
Answers vary

C.7 Contracts
Answers vary

C.9 No

0.1 (a) The field strength increases as the current increases.
(b) The field strength is inversely proportional to the distance.
(c) Yes
(d) When the current direction is away from you, the orientation of the magnetic field is clockwise

1.1 (a)
(d)

3.1 2.00e−5 T

3.3 1.14e−3 T

3.5 896 A

3.7 (a) 1.64e−26 N
(b) 55.4 m

3.9 4.58e−23 N, in the same direction as the current

3.11 (a) Opposite directions
(b) 75.0 A

3.13 8.80e−3 m

3.15 6.67 m

3.17 Direction: +*z*
Strength: 3.66e−7 T

3.19 (0, −1.60e−7, 1.20e−7) T

4.1 1.20 A, Left to right

5.1 (a) 3.05e−5 N/m, Attractive
(b) 3.05e−5 N/m, Repulsive

5.3 1.54e4 A

5.5 245 A

5.7 1.02e3 A

7.1 1.04 A

8.1 (a) 0 T
(b) 1.18e−6 T
(c) 2.00e−6 T

9.1 7.04e−4 A

9.3 1.99e−5 T

9.5 Yes
Answers vary

9.7 (a) 7.80e−6 N
(b) 0 N
(c) 0 N·m

10.1 8.00 A, Top to bottom

11.1 4.06e−2 m

11.3 (a) 3.89e−3 T
(b) 9.55e+5 turns/meter
(c) 3.24e−3 A

12.1 (a) 1.18e−2 T
(b) 7.65e−3 T
(c) 9.29e−3 T
(d) 0 T

12.3 (a) 1.02e−3 T
(b) 6.90e−4 T

Chapter 32 Answers

Chapter Assumptions

Unless stated otherwise, use $5.00\times10^{-5}\mathrm{T}$ for the strength of the Earth's magnetic field at its surface.

Elementary charge, $e=1.60\times10^{-19}\mathrm{C}$

Answers to selected problems

C.1 No
Answers vary

C.4 Clockwise
Answers vary

C.6 (a) Same as outer solenoid
Answers vary
(b) Same as outer solenoid
Answers vary

C.8 (a) Counterclockwise
(b) Clockwise
(c) Answers vary

C.10 (a) The same direction
Answers vary
(b) Answers vary

C.12 2

C.13 100

0.1 (a) No
(b) Yes
(c) Yes

2.1 (a) Yes
(b) The induced emf increases when the wire moves faster.
(c) The induced emf increases when the field strength increases.
(d) Answers vary

5.1 0.500 m

5.3 0.808 V

6.1 4.25 Wb

7.1 12.9 T· m^2

7.3 1.13e−6 A

8.1 8.00 ms

9.1 (a) 3.45e−8 W
(b) The power increases by a factor of 4
(c) Answers vary

13.1 289 m/s

13.3 (a) 2.5e3 W
(b) 1.9e−6 m^2
(c) 28 Ω

15.1 (a) The emf is a maximum when the velocity vector is Perpendicular to the magnetic field.
(b) The emf is zero when the velocity vector is Parallel to the magnetic field.
(c) Yes
(d) Yes

16.1 37.6 rad/s

16.3 1.26e+6 turns/m

17.1 1.80 m^2

20.1 1.50e+3 turns

20.3 26.2

20.5 (a) 7.75e3 for 100,000 V
1.74e4 for 225,000 V
2.67e4 for 345,000 V
(b) 900 A for 100,000 V
400 A for 225,000 V
261 A for 345,000 V
(c) 9.07e7 W for 100,000 V
1.79e7 W for 225,000 V
7.63e6 W for 345,000 V
(d) Answers vary

22.1 20 V

24.1 −34 A/s

24.3 2.17e−3 V

27.1 2.1e−5 H

29.1 2.93 A

30.1 (a) 6.22e5 J/m^3
(b) 6.91e−12 J/m^3
(c) 4

Chapter 33 Answers

Answers to selected problems

C.4 No
Answers vary

2.1 −5.7 A

2.3 (a) 623 rad/s
(b) 99.2 Hz

2.5 (a) 0 C
(b) 5.8e−4 C

5.1 0.021 s

6.1 (a) 0.2 A
(b) 2.0 V
(c) 1.0e6 Hz

9.1 (a) 1.39 A
(b) 1.39 A

10.1 $\pi/2$

11.1 8.49e−2 s

12.1 (a) 2.34 A
(b) 4.50 A

14.1 0.0413 s

15.1 (a) 31.0 A
(b) 16.1 A

16.1 (a) 1.27e−9 F
(b) 6.37e−5 H

17.1 (a) Capacitor
(b) Inductor
(c) Resistor

18.1 −0.246 rad

18.3 Capacitance corresponding to higher frequency: 1.73e−12 F
Capacitance corresponding to lower frequency: 1.80e−11 F

21.1 6.74e4 Ω

22.1 790 rad/s

22.3 1.29e6 rad/s

23.1 (a) 170 V
(b) -170 V

23.3 468 W

24.1 2.10 W

24.3 0.0229 W

26.1 8.80e−12 F

A.1 (a) 238 Hz
(b) 178 Ω
(c) 11.3 Ω
(d) 0.996 A
(e) −1.36 rad
(f) 17.4 W
(g) 238 Hz
(h) 238 Hz

A.3 (a) 5.0 V
(b) 0.0010 V
(c) 4.5 V
(d) 4.5 V
(e) 0.10 V
(f) 5.0 V
(g) The capacitor
(h) The inductor

Chapter Assumptions

Unless stated otherwise, electromagnetic waves are assumed to be propagating as plane waves through a vacuum.

When converting light-years to meters, assume that a year has 365 days.

The speed of light in a vacuum is 3.00×10^{8}m/s.

The Sun radiates energy at the rate of 3.91×10^{26}W.

Answers to selected problems

C.1 AM
Answers vary

C.3 Yes
Answers vary

C.5 (a) North−south
(b) Vertical

C.7 (a) Vertical
(b) Vertical

1.1 (a) 1.28 light−seconds
(b) 8.34 light−minutes
(c) 2.6e20 m

1.3 (a) 7.5e5 Hz
(b) AM

1.5 (a) 12.2 cm
(b) 6.10 cm
(c) Yes

1.7 (a) (0, 0, 0)
(b) (255, 255, 255)
(c) 16777216 colors

1.9 (a) 19%
(b) 57%
(c) 16%

2.1 Away from you

3.1 1.39e−6 T

7.1 (a) 272 W/m^2
(b) 136 W/m^2
(c) 136 W/m^2
(d) 226 V/m

7.3 (a) 1.59e−2 J/m^3
(b) 1.59e−2 J/m^3
(c) 3.18e−2 J/m^3

7.5 1.25e5 J

7.7 (a) 1.83e−5 J/m^3
(b) 2.33e−3 A

10.1 8.55%

10.3 4.12e−7 W/m^2

10.5 7.78e−25 W/m^2

10.7 (a) 1.2e−8 T
(b) 1.6e−2 W/m^2
(c) 2.0e−5 W

12.1 (a) 4.0e6 m
(b) 2.0e6 m
(c) 1.0e6 m

13.1 (a) 2.28e−2 N
(b) The weight of the coins

13.3 (a) 3.98e−8 N/m^2
(b) 7.96e−8 N/m^2

13.5 (a) 8.15e8 N
(b) 3.52e22 N

13.7 2.06e−7 m

18.1 106 V/m

18.3 97.3%

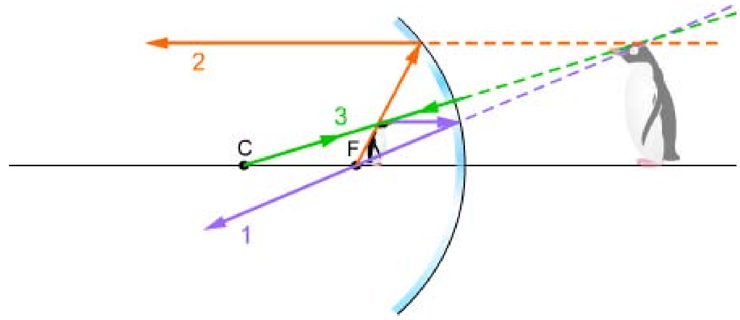
22.1 2.504 °m^2/kg

Chapter 35 Answers

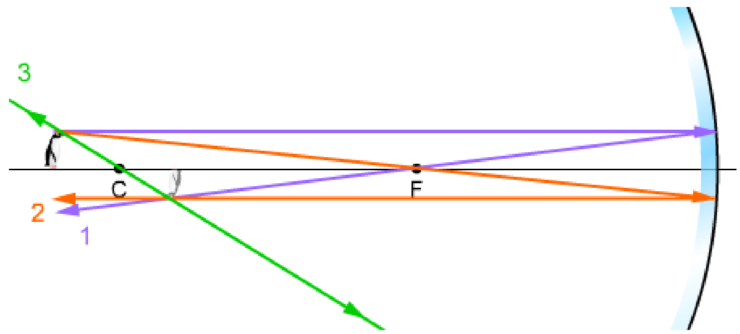
Answers to selected problems

C.1 (a) 3
(b) B

C.3 Virtual
Larger
Upright



C.5 Real
Smaller
Inverted



C.7 (a) Larger
(b) Inverts
Becomes real

C.9 Convex
Answers vary

0.1 (a) To the left of C
In between C and F
(b) To the left of C
(c) In between C and F
To the right of F

3.1 7.50 m

5.1 (a) 1.19 m
(b) 53.5°
(c) 180°

5.3 22

5.5 1.56 m

7.1 (a) 0.600 m
(b) 0.570 m

10.1 -1.15 cm

10.3 10.5 cm

15.1 (a) No
(b) No
(c) No
(d) No

16.1 (a) +15.2 cm
(b) −0.879
(c) Inverted

16.3 16.5 cm

16.5 (a) 12.5 cm
(b) 8.33 cm

16.7 63.3 cm

16.9 (a) Concave
(b) 6.20 cm

16.11 −34.0 cm

16.13 (a) Inverted
(b) Virtual
(c) 29.6 cm

18.1 6.16 cm

Chapter 36 Answers

Chapter Assumptions

Unless stated otherwise, the incident medium is air, with an index of refraction of 1.00.

Answers to selected problems

C.1 (a) X
(b) Y

0.1 (a) Air
(b) Toward

2.1 Diamond

2.3 1.86e8 m/s

3.1 12.0°

3.3 1.27

3.5 1.63

3.7 1.71

7.1 1.92

10.1 (a) 65.8°
(b) 34.8°

11.1 24.4°

11.3 1.54

13.1 (a) 0.045 m
(b) 0.330 m

14.1 0.11°

14.3 (a) 39.6°
(b) 31.5°
(c) 44.5°

A.1 3.84 m

A.3 29°

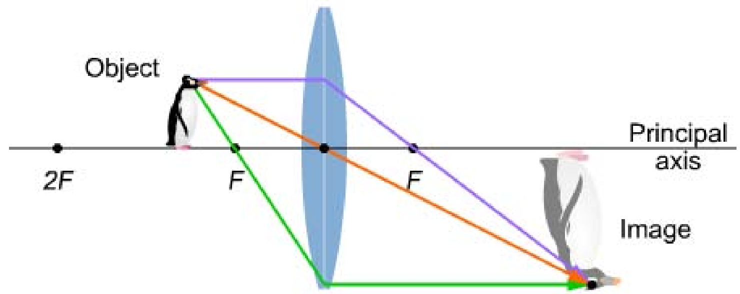
A.5 1.50

A.7 1.89 m

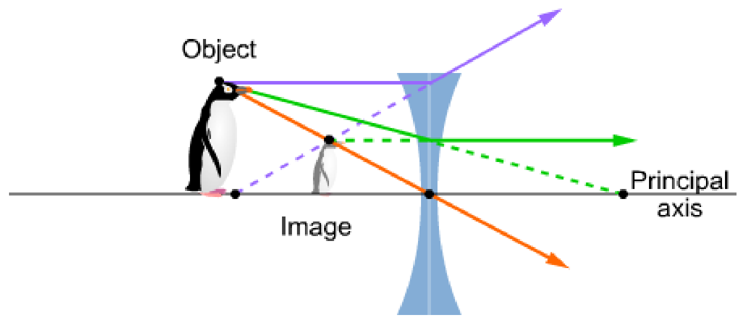
Chapter 37 Answers

Answers to selected problems

C.1 Real; larger; inverted



C.3 Virtual; smaller; upright



C.5 Inverted
Answers vary

C.7 Speed of a bullet

C.9 (a) At the focal point
(b) Answers vary

0.1 (a) Between the focal point and the lens.
(b) On the opposite side of the focal point as the lens.
(c) Between the focal point and the lens.

4.1 (a) No
(b) No
(c) No

7.1 (a) Positive
(b) Positive
(c) 4.11 cm
(d) Converging

7.3 0.135 m

7.5 (a) Negative
(b) Real
(c) 14.0 cm
(d) 4.67 cm

7.7 50.0 cm

7.9 (a) Diverging
(b) −3.94 cm

7.11 0.0233 m

7.13 1.18 m

9.1 4.03 cm

10.1 (a) 10.0 cm
(b) 6.00 cm

14.1 (a) Diverging
(b) −40.0 cm

17.1 1.54 diopters

17.3 −0.400 m

18.1 0.0066 radians

18.3 1.40e5 m

19.1 8.40 cm

23.1 (a) -53
(b) Decrease
(c) −28

23.3 (a) −34.4
(b) 7.67e−3 radians

24.1 −9.0e2

A.1 -6.65 diopters

A.3 17.3 cm

A.5 (a) At infinity
(b) Virtual
(c) 40.0 cm

A.9 0.246 cm

A.11 (a) 1.16 cm
(b) 0.255 m^2
(c) 2.41e6 W/m^2

Chapter 38 Answers

Chapter Assumptions

Unless stated otherwise, use the following indices of refraction

n_{air} = 1.00

n_{water} = 1.33

n_{glass} = 1.50

Answers to selected problems

C.1 $\lambda/2$

C.5 An odd number of half−wavelengths

C.7 The outlying fringes move closer to the centerline.

C.11 Blue
Answers vary

3.1 (a) 560e−9 m
(b) 6.7e−3 m

3.3 6.96e−2 m

3.5 581e−9 m

3.7 0.145 °

3.9 4.73e+14 Hz

4.1 2.16e−3 m

4.3 805 nm

6.1 606e−9 m

6.3 68 fringes

9.1 (a) $2nt = (m + 1/2)\lambda$
(b) $2nt = m\lambda$

9.3 1.21e−7 m

9.5 (a) Less
(b) 664e−9 m

11.1 161 bright fringes

11.3 15 fringes/cm

12.1 6.16e-7 m

A.1 3.11e−7 m

A.3 4.58e−7 m

Chapter 39 Answers

Answers to selected problems

C.1 No
Answers vary

C.3 Red light
Answers vary

5.1 (a) 7.46e−3 m
(b) 1.49e−2 m

5.3 3.04e−2 m

5.5 524 nm

5.7 8.07e−5 m

5.9 7.41e−6 m

7.1 1.05e−2 m

11.1 1.4 m

11.3 (a) 5.39e4 m
(b) 50.7 m

13.1 1.02e−6 m

13.3 0.633°

A.1 6.00e−5 m

Chapter 40 Answers

Chapter Assumptions

Assume the speed of light in a vacuum is 3.00×10^8 m/s for all observers.

Answers to selected problems

C.1 No

C.3 (a) 1.00*c*
(b) 1.00*c*

C.5 No

C.7 *V*/2

3.1 (a) 9.0 s
(b) 8.5 s

7.1 (a) Yes
(b) Yes
(c) Yes
(d) No

8.1 (a) 3.00e−7 s
(b) 3.25e−7 s

8.3 1.55e+8 m/s

8.5 2.89e+8 m/s

10.1 (a) 555 m
(b) 745 m
(c) 1.85e−6 s
(d) 2.48e−6 s
(e) 1.34
(f) 2.48e−6 s

11.1 (a) 6.0e−8 s
(b) 1.0e−7 s
(c) Shorter

12.1 (a) 0.333*c*
(b) 65.0 m

13.1 0.948*c*

13.4 (a) 198 m
(b) 1.13e−6 s

14.1 (a) 2.72e8 Hz
(b) 2.72e8 Hz

14.3 (a) 1.20 s
(b) 1.27 s
(c) 1.69 s

14.5 0.111*c*

16.1 (a) 0.875*c*
(b) 1700 MeV/c

16.3 1.32 MeV

16.5 (a) 0.163 km
(b) 1.46e+13 years

16.7 5.9e4

A.2 (a) 7.83 m
(b) 36.4°

Chapter 41 Answers

Chapter Assumptions

When converting between electron-volts and joules, assume that

1 eV = 1.60×10^{-19} J

Answers to selected problems

C.3 (a) There are more electrons leaving the surface.
(b) The electrons' maximum kinetic energy is still equal to E_0.

C.5 There is movement of holes in the valence band.

C.9 (a) Positive
(a) Negative

C.11 Yes
Answers vary

2.1 486 nm

4.1 4.2e14 Hz up to 7.5e14 Hz

4.3 1000000

6.1 4.74 eV

6.3 12 nm

9.1 shorter wavelength: 122 nm
longer wavelength: 658 nm

13.1 1.11e−6 m

16.1 2.4 eV

Chapter 42 Answers

Chapter Assumptions

1 eV = 1.60×10^{-19} J

$h = 6.63\times10^{-34}$ J·s

$\hbar = 1.05\times10^{-34}$ J·s

$m_e c^2 = 0.511$ MeV

Answers to selected problems

C.1 Electron A
Answers vary

C.3 Becomes smaller
Answers vary

1.1 1.19e−27 kg·m/s

2.1 7.81e−11 m

2.5 (a) 0.18 MeV
(b) 6.9E−3 nm
(c) 170°

2.7 159 keV

4.1 2.01e−38 m

4.3 1.67E−27 kg

4.5 120 V

4.7 2.65e3 MeV

6.1 (a) 1.41e−10 m
(b) 2.80e3 m/s

10.1 (a) 5.0e−6 m
(b) 1.1E−29 kg·m/s

10.3 (a) 3.3e−20 m
(b) No

Chapter 43 Answers

Answers to selected problems

C.1 A small fraction were deflected from the gold foil.

C.3 (a) ^{12}C and ^{14}C
^{14}O and ^{16}O
(b) ^{12}C and ^{14}O
^{14}C and ^{16}O
(c) ^{12}C and ^{14}C
^{14}O and ^{16}O

C.5 Electric (Coulomb) attraction

C.7 No
Answers vary

C.9 Nuclear fusion

3.1 19 electrons

3.3 (a) 40 is the mass number
(b) 20 protons
(c) 20 neutrons
(d) 18 electrons

6.1 64

7.1 1.8e2 m

8.1 206

9.1 4.54e−12 J

13.1 2.96e−11 J

14.1 (a) 4 hydrogen atoms
(b) 4.62e−29 kg

14.3 1.3e11 s

15.1 (a) Decrease
(b) 1.3e−29 kg

15.3 One alpha and two negative beta particle(s)

15.5 Lead−206

16.1 Negative beta particle

16.3 (a) Yes
(b) 82
(c) 205

18.1 6.25 percent

19.1 1.42e14 cobalt−60 atoms

19.3 4.70 days

21.1 5730 years