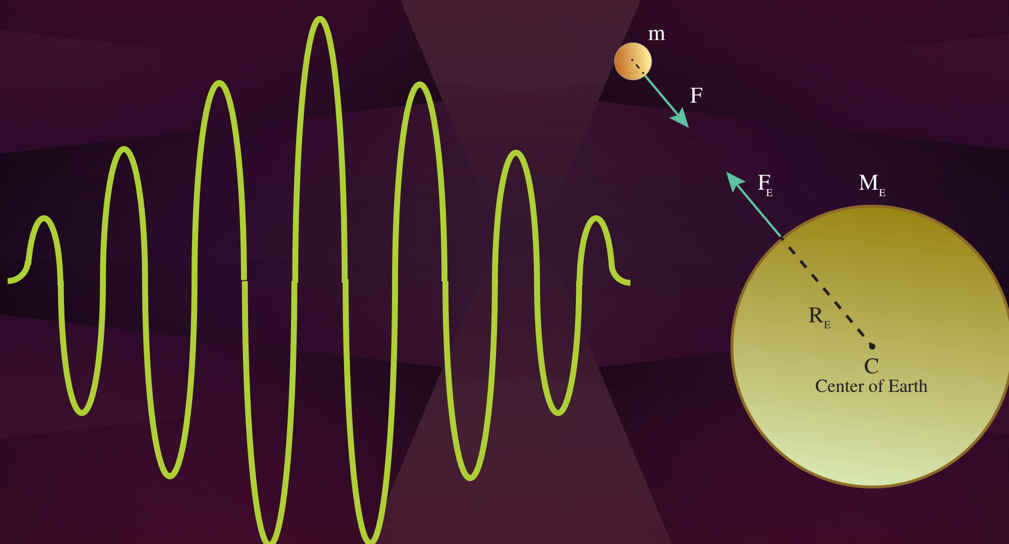
ESSENTIAL PHYSICS



John Matolyak
Ajawad Haija

CRC Press
Taylor & Francis Group

ESSENTIAL PHYSICS

ESSENTIAL PHYSICS

John Matolyak
Emeritus Professor, Indiana University of Pennsylvania

Ajawad Haija
Indiana University of Pennsylvania

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20131101

International Standard Book Number-13: 978-1-4665-7524-0 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedicated to my late mother, Anna (Radak) Matolyak

Her constant love, encouragement, and support made much possible.

John Matolyak

Dedicated to my wife, Samira Hassan, whose support and sacrifices have been beyond any level of gratitude.

Ajawad Haija

Contents

Preface

Many university degree programs require their students to complete a 1-year course in noncalculus general physics. The purpose is to give students an understanding of the basic laws of physics. These laws are part of the foundation of the student's chosen field of study, which is usually not physics. Physics, in large part, is about fundamentals. More recent general physics texts have these fundamentals interspersed between many applications and special topics. Additionally, these texts have typically between 40 and 90 problems at the end of each chapter. Inclusion of all these materials results in a text of the order of 1000 pages.

It is our experience that in a 14- to 17-week semester, two-semester course, much of the text material must be omitted and only a small number of chapter problems can be assigned. Much of the large and expensive text remains unused.

This work is an attempt to emphasize the fundamentals of physics and gloss over, as unpleasant as it is for us, the other "goodies" of physics. The title *Essential Physics* was chosen to emphasize the need for nonphysics majors to understand the physics essential to their major fields of study. For instance, just a few of the physics essentials that chemistry majors need to be familiar with are electric fields and electrostatics in order to understand the structure of atoms and molecules. Biology majors need to understand fluid flow, muscle forces, heat loss by respiration, and so on. Safety science majors should be familiar with conservation of linear momentum as applied to collisions. These majors do not need to understand the microscopic physics of a transistor, how a television picture is formed, or the details of an inkjet printer.

We think that fundamentals are more readily learned and mastered by student exposure to a combination of worked example problems and narrative discussion of theory. An attempt has been made to balance the two.

Some fluency with the fundamentals of physics and physics problem solving has a collateral effect on the student. It enhances his or her analytical reasoning skills. In a sense, physics is to other intellectual pursuits what weightlifting is to other sports. It is quite appealing in its own pursuit and very "strengthening" and useful in the pursuit of others.

This book is appropriate for a course in which the goals are to give students a grasp of physics fundamentals and enhance their analytical problem-solving skills. If exposure to more applications, special topics, and concepts is desired, this book can be used as a problem-solving supplement to a more inclusive text.

Authors

John Matolyak was initially employed as a lab technician researching the purification of silicon at Union Carbide Laboratory in Parma, Ohio. After acquiring an undergraduate degree (double majors: mathematics and physics), he worked for NASA (Goddard) investigating the back-scattering of laser light from the atmosphere. He earned his PhD in 1975, and his dissertation addressed antiferromagnetic magnetostriction. In addition, he worked at NASA (Lewis) and the U.S. Army's Night Vision Laboratory.

He has taught and directed research in physics at both the undergraduate and graduate levels at three separate universities. He is currently an Emeritus Physics Professor at the Indiana University of Pennsylvania.

Ajawad Haija, born in Palestine, grew up in Jordan and attended the University of Alexandria in Egypt, where he received his BSc degree in 1968 with distinction, 1st honor. He received his PhD (1971–1977) at Pennsylvania State University in 1977. His dissertation in solid-state physics addressed electrical and optical properties of super-lattices. Dr. Haija joined the University of Jordan in Amman, Jordan, in 1977, where he served until 2000. In 2000, he decided to move to the United States and joined the Indiana University of Pennsylvania. In addition to numerous publications in the field of thin films, multilayers, and super-lattices, he has authored two books, both in Arabic: *Classical Physics*, published in 1998, and another (in literature), *Satirical Thoughts from the Inspiration of the Twentieth Century*, published in 2000. He has also translated into Arabic, reviewed, and edited numerous translations of the Arabic version of *Scientific American*, which is published in Kuwait.

He is currently on the faculty of Indiana University of Pennsylvania, professor of physics, where he conducts his research on the optical properties of thin multilayer structures and super-lattices. He enjoys teaching physics courses that extend from freshman to senior-level students. He is a member of the American Physical Society.

1 Systems of Units, Significant Figures, Coordinate Systems, and Vectors

Measurements play an essential part in physics, and the purpose of measurement is to acquire a number or numbers that accurately represent the values of the measured physical quantities under consideration. However, it would be impossible to relate the experimentally obtained numbers to their physical quantities without defining a standard, to which they could be compared and then documented. For example, it would be insignificant to say that a certain quantity has a mass of 300 g without having in mind a standard of mass for comparison. The standard of mass of 1.000 kg makes the comparison so clear because we understand then that 300 g are 300 parts of 1000 parts that make 1 kg. The same applies to length and time.

1.1 SYSTEMS OF UNITS

It is not too hard to anticipate the possible existence of several systems of units that evolved with man's quest for measurements and meaningful analyses that followed these measurements. In fact, there are three systems of units that are commonly used.

1.1.1 The cgs System

The cgs system adopts the centimeter (cm) as a unit of length, the gram (g) for mass, and the second (s) for time.

1.1.2 The mks System

The mks system is also known as the SI system. This system at present is the most scientifically recognized and widely used, and it adopts the meter (m) for length, the kilogram (kg) for mass, and the second (s) for time.

1.1.3 The ft–lb–s System

The ft–lb–s system is also known as the BE (British Engineering System). It adopts the foot (ft) for length, the pound (lb) for weight, and the second (s) for time. This system was in common use in Britain and in some of the previously British-influenced regions. We can realize as we go on that the first and the third systems are not too popular in the sciences, and the interest has always been in adopting the mks system. In this chapter, we will adhere as often as possible to the mks system.

With regard to units and systems of units, the following are the two points to note:

1. Although each of the systems described earlier is comprised of only three fundamental quantities, one encounters numerous physical quantities that need to be described by other units more involved than just one kind of unit. In this case, a derived unit that is comprised of more of the fundamental quantities is used. Examples of this kind are the area whose

units are m^2, velocity whose units are m/s, and mass density, kg/m^3. Physical quantities of derived units will be encountered frequently in this chapter.

2. As a need arises for executing some calculations that involve two or more physical quantities, care must be taken that all physical quantities have to be described in any one of the systems of units, the mks, the cgs, or the ft–lb–s system. A mix of units that belong to two different systems is not allowed. If a problem happened to present quantities described by units from different systems such as the cgs and the mks, then a conversion of those cgs-described quantities to their corresponding values in the mks has to be carried out prior to executing the desired calculation.

1.2 CONVERSION OF UNITS

Since there are more than one system of units and numerous physical quantities of various magnitudes, huge or small, a large number of units were devised to describe those physical quantities. Accordingly, conversion of units is inevitable. Table 1.1 lays out in powers of 10 many of the prefixes that are frequently used in different fields of physics and that can be useful for units conversion purposes and in involved calculations.

EXAMPLE 1.1

A square tile is 30.48 cm on a side. Calculate its area in m^2.

Solution

As the area of the square tile is $A = (\text{side})^2$, then this in m^2 is

$$A = (30.48 \times 10^{-2}\ \text{m})^2 = 0.09290\ \text{m}^2.$$

To have a feel for this number, one can determine the number N of tiles that would be accommodated in 1.00 m^2, we just divide 1.00 m^2 by the area of each tile. Thus,

$$N = \left(\frac{1.00\ \text{m}^2}{0.0920\ \text{m}^2}\right) = 10.8.$$

The given 30.48 cm for each side of the tile is in fact 1.000 ft. Therefore, the area of the tile is 1.000 ft^2. Accordingly, slightly <11 tiles can fit in 1.000 m^2.

TABLE 1.1
Prefixes That Express Subdivisions and Multiples of Units in the mks System

Power of 10	Symbol	Abbreviation
10^{-15}	f	Femto
10^{-12}	p	Pico
10^{-9}	n	Nano
10^{-6}	μ	Micro
10^{-3}	m	Milli
10^{-2}	cm	Centi
10^{3}	k	Kilo
10^{6}	M	Mega
10^{9}	G	Giga
10^{12}	T	Tera

EXAMPLE 1.2

The density of pure water is $\rho = 1.00\ \text{g/cm}^3$. Determine this density in kg/m^3.

Solution

Using Table 1.1 to express the gram and centimeter in their mks units, the density of water becomes

$$\rho = \left(1.00\,\text{g} \times 10^{-3}\,\frac{\text{kg}}{\text{g}}\right) \Big/ \left(1.00\,\text{cm} \times 10^{-2}\,\text{m/cm}\right)^3 = 1.00 \times 10^3\,\text{kg/m}^3$$

As one can see, the weight of 1.00 cubic meter of water is 1000 kg, that is, there is one metric ton for each cubic meter. This is a rather large value that gives you a clear idea of the possible impact that occurs to disastrously flooded areas.

1.3 SIGNIFICANT FIGURES

The importance of significant figures originates from the fact that measuring tools and devices are limited in precision according to specifications of their design. For example, if a balance scale has the gram as its lowest division, tenths of a gram could only be an estimate on that scale. In reporting any number, we are allowed to keep only one doubtful digit among all the digits of that number. Therefore, any reported mass weighed by this scale is reliable as long as the mass is reported in grams. If the documented mass had a digit that indicates tenths of a gram, then that last digit, although allowed, is not certain, and definitely any further precision claimed by reporting the number with digits that correspond to hundredths of a gram is meaningless.

For illustration of this case, let us take a numerical example as 5.4 g. The number of the significant figures in this value, 5.4, is two, and the digit 4 is doubtful. We cannot report this number as 5.41, because it would mislead concerned individuals of an unexisting precision, which the balance did not have. So the rule is to have only one digit that is uncertain. The allowed uncertain digit here is 4, and we should terminate the number at the digit 4. The same principle applies to a ruler that has a scale in centimeters and millimeters. So, a reading that includes millimeters should be precise and estimation of a number that is read on this ruler is allowed to only one digit beyond the millimeter. Therefore, a reading of 5.72 cm means that the reading up to 5.7 cm is certain and the digit 2 is uncertain, which we are allowed to report. However, any digit beyond that is meaningless and is not permitted; otherwise, a false precision is claimed. The number of the significant figures in the quantity 5.72 is three.

1.3.1 Determination of the Number of Significant Figures

The issue of significant figures stems from the fact that measuring tools and devices are limited in precision according to specifications of their designs. A quantity that is either measured or calculated should have as many digits as needed to recognize the measuring instrument's precision. The following rules can be very helpful for recognizing the significant figures:

1. All nonzero digits in a certain number are considered significant figures.
2. The zero digits that lie between nonzero digits in a certain number are also considered significant figures.
3. The zero digits that lie to the right of the last nonzero right digit with no decimal point are not counted as significant.
4. In a decimally expressed number, the zero digits that lie to the right of the last nonzero digit are counted as significant figures.
5. In a decimally expressed number that is <1, the zeros between the decimal point and the first none zero digit are not counted as significant.

1.3.2 Rounding Off

1. As we intend to omit some digits from those that lie to the right of a decimal point, we check the digit following the number to be retained. If that digit is below 5, then that digit and all of those to its right can be dropped leaving the remaining digits unaltered, that is, we round "down."
2. If in a certain number the first digit is to be dropped, following the digit to be retained is 5 or higher, then we add 1 to the digit on its left. In other words, the extreme right digit that will survive will become higher by 1 than what it was before dropping those following on its right, that is, we round "up."

EXAMPLE 1.3

Round off the following number 25.2463 such that it has

a. Only five significant figures
b. Only four significant figures

Answer and Analysis

a. Rounding off 25.2463 to only five significant figures means that the digit 3 to the right of the decimal point should be dropped. As we do so, the finally accepted number would be 25.246.
b. Rounding off 25.2463 to only four significant figures means that the digits 63 to the right of the decimal point should be dropped. But since the first digit to drop is higher than 5, we add 1 to the first digit retained on the left of the omitted 6. The finally accepted number would then be 25.25.

1.3.3 Significant Figures in Arithmetic Operations

Upon executing some basic mathematical operations, addition, subtraction, multiplication, and division, an occasional confusion may rise in regard to the number of the significant figures, which an answer resulting from such operations should contain. This is because executing any of those operations, multiplication and division in particular, may have more significant figures than what is relevant to the quantity being calculated. The best way to explain is through several examples, one at least on each operation.

1.3.4 Addition and Subtraction

Upon adding or subtracting some numbers, the resulting answer should have as many digits after the decimal point as the number of digits after the decimal in the least precise number.

EXAMPLE 1.4

Perform the operation, 25.13 cm + 1.7 cm + 0.0054 cm, keeping the proper number of significant figures.

Answer and Analysis

The answer as executed on a calculator would be 26.8354 cm. However, 1.7 is known only to a tenth of a centimeter. Therefore, the sum can only be known to a tenth of a centimeter. The final answer is 26.8 cm.

EXAMPLE 1.5

Perform the operation, 25.71 cm – 1.2 cm, keeping the proper number of significant figures.

Answer and Analysis

The concept applied in Example 1.4 applies here to the subtraction process as well. The answer 24.51 cm as executed on a calculator cannot be more precise than the least precise number that entered into the subtraction. Therefore, dropping the digit 1 should not invalidate the precision of the subtraction operation. Accordingly, the final answer would be 24.5 cm.

Thus, in addition and subtraction, one must attend to the proper number of decimal places, not the significant digits. That is, the final sum or difference cannot be more precise than the least precise of the terms added or subtracted.

1.3.5 Multiplication and Division

In the quantities that enter into a multiplication process, the obtained result should have a number of significant figures equal to that in the least precise number of those involved in multiplication. In other words, the decisive factor regarding the number of the significant figures in the end result of a multiplication or division operation is the lowest number of significant figures in the quantities that formed that operation. The number of the significant figures in the result of the operation should not exceed the number of the significant figures in the least precise quantity. The following examples can illustrate this process.

EXAMPLE 1.6

In the following, let $\rho = 0.03$ g/cm^3 be the density of some very light alcohol derivative with volume $v = 0.0021$ cm^3. Determine the mass, m ($m = \rho \times v$), of this quantity.

Answer and Analysis

The computed product would be 0.000063 g. In the quantities that entered into the multiplication process, we had only one significant figure in the number 0.03, while the other, the 0.0021, had two significant figures. Since the obtained result is not expected to be more precise than the quantity that had the least number of significant figures in the multiplication operation, the final answer should be satisfactory enough if we report it as a product that has only one significant figure, 0.00006. That is, in a round "down," the digit 3 that lies to the right of the digit 6 was omitted in the final answer.

EXAMPLE 1.7

Determine the area of a piece of land of length, $L = 71$ m, and width, $W = 15.7$ m.

Answer and Analysis

The product of these two numbers is 1114.7 m^2, but since one of the numbers, 71, being the lowest in the number of significant figures, the product should have only two significant figures. Therefore, we drop in the product all digits starting with the third digit and beyond. The resulting product then should be 1100 m^2.

EXAMPLE 1.8

In the following, a 4 m^3 of some dirt is to be spread over a lawn of 6 m^2 area. Assuming that the spread is uniform, determine the thickness of the dirt spread over the whole area.

Answer and Analysis

The division of the given volume by the specified area gives the required thickness, which on a calculator is 0.66666 m, but for this answer, one should not have more than one significant figure. Therefore, with rounding the answer up, one would obtain that as 0.7 m.

1.4 SCIENTIFIC NOTATION

Scientific notation provides a convenient way of expressing numbers that are either very small or very large. Typical examples in physics, especially in areas of astronomical and atomic dimensions, are numerous. A highly convenient way to report such quantities is in powers of 10. The sign of the exponent of 10 would be positive if the quantity under consideration is huge and negative if it is rather small. The rule goes as follows. As 10 is simply 10^1, 100 is 10^2, and 1000 is 10^3, then quantities such as 549 and 5492 can be written as 5.49×10^2 and 5.492×10^3, respectively. As can be seen, the number 549 has been expressed as the product of 5.49 times 10 raised to a power of 2. The number of the digits to the right of the decimal point is equal to the exponent of 10 that is introduced. The same rule is applied to 5492. In brief, the number is written in a decimal form between 1 and 10 times some power of 10 that is determined by how many digits are kept to the right of the decimal. For small numbers, however, the following examples illustrate a similar method, except that the exponent is negative. As $0.1 = (1/10) = 10^{-1}$, $0.01 = (1/100) = 10^{-2}$, $0.001 = (1/1000) = 10^{-3}$, and so forth, then 0.147 and 0.0147 can be written as 1.47×10^{-1} and 1.47×10^{-2}, respectively. The following examples offer further demonstration of this issue.

EXAMPLE 1.9

The speed of light in vacuum c is 299,792,458 m/s (E. R. Cohen and B. N. Taylor, *Journal of Research of the National Bureau of Standards*, 92, 85, 1987). Write this number in scientific notation.

SOLUTION

Following the above rule, the speed of light in vacuum, c, in scientific notation becomes

$$c = 2.99792458 \times 10^8 \text{ m/s},$$

which is usually rounded off to two significant figures 3.0×10^8 m/s.

EXAMPLE 1.10

Write down the number 0.0012 in a scientific notation.

SOLUTION

The number 0.0012 in scientific notation becomes 1.2×10^{-3}.

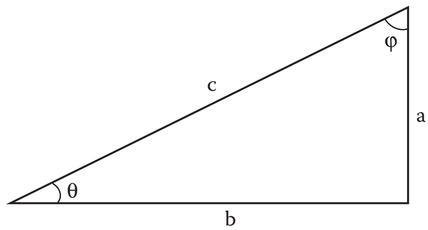
1.5 TRIGONOMETRY

Geometrical definitions: In the right angle triangle (the following figure), the following trigonometric relations are defined:

$$\sin\theta = \frac{a}{c}, \tag{1.1}$$

$$\cos\theta = \frac{b}{c}, \tag{1.2}$$

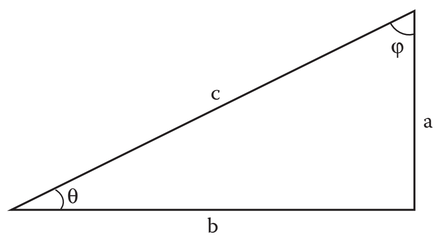
$$\tan\theta = \frac{a}{b}, \tag{1.3}$$



EXAMPLE 1.11

A right triangle (the figure below) has the sides a = 2.50 m and b = 6.00 m. Calculate

a. The angles θ and φ
b. The hypotenuse c



Solution

a. From Equation 1.3,

$$\tan\theta = \frac{a}{b},$$

which after substituting for a and b becomes

$$\tan\theta = \frac{2.50}{6.00}$$

or

$$\tan\theta = 0.417.$$

Thus,

$$\theta = \tan^{-1}(0.417) = 22.6°.$$

Similarly,

$$\tan\varphi = \frac{b}{a} = \frac{6.00}{2.50} = 2.40.$$

Thus,

$$\varphi = \tan^{-1}(2.40) = 67.4°.$$

Analysis

One can see that the two angles θ and φ add up to 90°, which is what is expected for the two small angles in a right angle triangle.

b. From Equation 1.1, we have

$$\sin\theta = \frac{a}{c}.$$

Substituting for θ, we then have

$$\sin 22.6° = \frac{(2.5)}{c}.$$

That is,

$$0.38 = \frac{2.5}{c}.$$

Solving for c gives

$$c = \frac{2.5}{0.38} = 6.6\text{ m}.$$

Analysis

One can verify Pythagorean theorem that states that $c = \sqrt{a^2 + b^2}$, and hence c could have been obtained using the Pythagorean theorem in place of using the trigonometric relation.

1.6 COORDINATE SYSTEMS

Coordinate systems in various disciplines, physics, pure and applied mathematics, geometry, economics, financial data, and engineering, are very useful to display curves, charts, and variations of related quantities. In physics, kinematics in particular, coordinate systems are of a special and frequent use for representing locations of objects in space and time. To demonstrate a simple and common use of coordinate systems, we will consider coordinate systems in two dimensions and, accordingly, treat objects in only a two-dimensional space. Cartesian coordinate systems will be given special attention since they facilitate much of our understanding of numerous fundamental quantities that occur in kinematics and dynamics. Among such quantities are displacement, velocity, acceleration, force, momentum, field, and torque, each of which is described by two parameters, magnitude and direction. These quantities are called vectors and obey special rules and mathematical operations. As one will recognize later in this chapter, Cartesian coordinate systems are a powerful tool to articulate these required mathematical and geometrical operations.

Among the many uses of coordinate systems is their application to the description of an object location in space. There are several coordinate systems among which the following two are well known:

1. The Cartesian coordinate system that consists of two perpendicular axes, usually labeled as x and y, and a reference point O, defined by the intersection of the two axes (Figure 1.1). The coordinates of the point P then are (x, y).
2. The plane polar coordinate system that consists of a coordinate r measured from an origin O, along a line that makes an angle θ with a reference line that starts from the origin and extends in a specific direction, usually the horizontal direction (Figure 1.2). The coordinates of the point P then are (r, θ), where r is the coordinate, that is the distance, along a line that makes an angle θ with the reference line. In the following, a more detailed discussion on these two systems is presented.

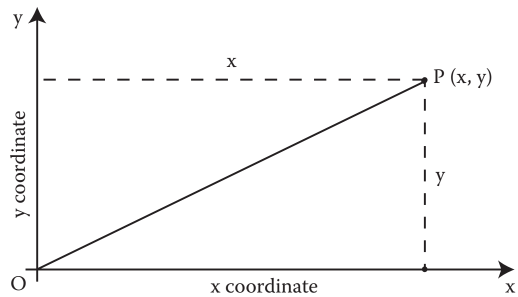


FIGURE 1.1 A two-dimensional Cartesian coordinate system; x and y are the axes and (x, y) are the coordinates of the point P.

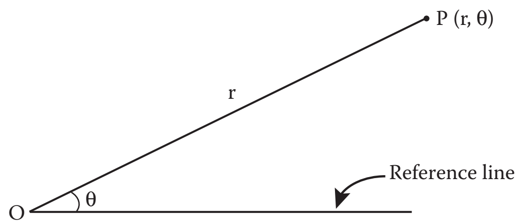


FIGURE 1.2 A plane polar coordinate system; r and θ are the two coordinates of the point P.

1.6.1 Cartesian Coordinate Systems

We construct a two-dimensional Cartesian coordinate system as follows:

1. Draw two perpendicular lines that we adopt as two axes, x and y.
2. Consider the intersection of the above two axes as a reference point, O, and adopt it as the origin to which all measured locations are referred (Figure 1.3). The coordinates of the origin, O, are (0, 0).

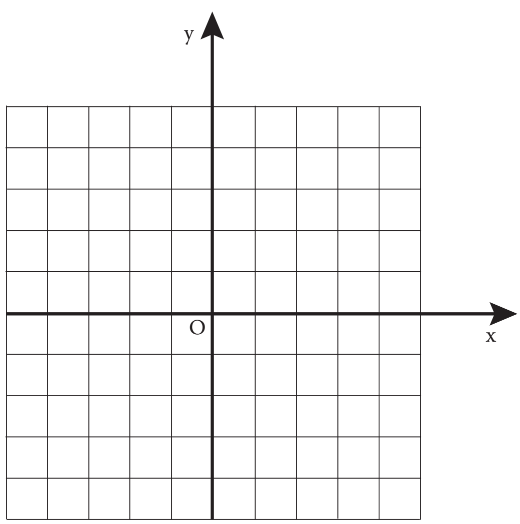


FIGURE 1.3 A Cartesian coordinate system consisting of two perpendicular axes, x and y, each of which extends from negative infinity to positive infinity. The axes intersection defines the origin O (0, 0).

The arrow on each axis indicates the positive direction along that axis. The light horizontal and vertical lines indicate divisions along the two axes according to which a certain coordinate is counted. Starting from the origin whose coordinates are (0, 0), we move along each axis by the number of units that define the coordinate of the position along that axis.

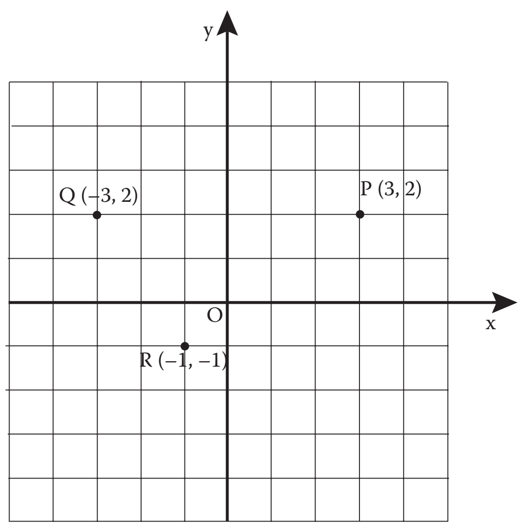
EXAMPLE 1.12

Locate on a Cartesian coordinate system the points P, Q, and R, whose coordinates are (3, 2), (–3, 2), and (–1, –1), respectively.

Answer and Analysis

With the divisions displayed on the axes of the following figure, a point P whose coordinates are (3, 2) is positioned by moving away from the origin, O, along the x-axis 3 divisions (units), and moving along the y-axis 2 divisions (units).

In a similar way, the other points Q (–3, 2) and R (–1, –1) can easily be located.



1.6.2 Plane Polar Coordinate Systems

The plane polar coordinate system consists of two coordinates, r and θ. A two-dimensional polar coordinate system is constructed as follows:

1. Starting from a point that may be considered an origin O, draw a line that is adopted as a reference line, usually sketched along the horizontal direction. This line may also be called the zero angle line.
2. Sketch another line that extends from O and makes an angle θ with the reference line (Figure 1.4). Along this new line, locate the point P at a distance r from O. The coordinates of the point P then are (r, θ). The angle θ is measured counterclockwise.

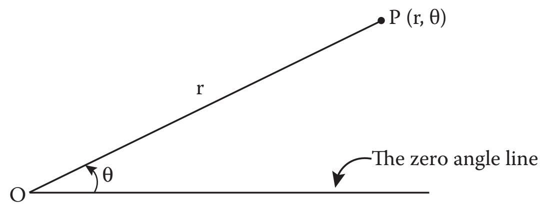


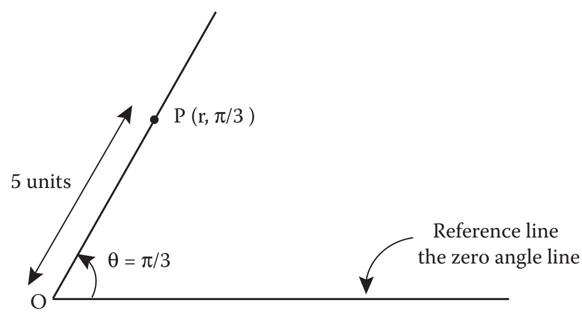
FIGURE 1.4 The figure shows a point P of plane polar coordinates (r, θ). The reference line is the zero angle line, and OP (= r) is a line that starts from O and makes an angle θ with the reference line.

EXAMPLE 1.13

Locate on a plane polar coordinate system a point P whose coordinates are (5, π/3).

Answer and Analysis

Starting from a point, O, first draw a reference line, and from O, then draw another line OP that makes an angle $\theta = \pi/3$ with the reference line. Along this new line, locate a point P that is away from O with a distance of 5 units (the figure below). The point P then is the point of coordinates (5, π/3).



1.7 VECTORS

Two classes of physical quantities that are of interest are vectors and scalars. A vector is a quantity that has a magnitude and direction. Examples of vectors are displacement, velocity, acceleration, force, and so many other quantities that we will encounter in our way through this chapter. A scalar is a quantity that can have a magnitude only. Examples of scalars are length, volume, mass, mass density, and time.

A physical vector quantity cannot be fully described without a preknowledge of its magnitude and in what direction the quantity is pointing. For example, the velocity of a car is by definition a vector and has no relevance without stating the value of the car's velocity and the direction along which the car is heading, although the speed of a car is a scalar. The same is true for all vectors.

What are the mathematical properties of vectors? The main mathematical operations that are discussed here are simply addition and subtraction. This is because most of the applications, handled here, will be confined to determining the resultant of two or more vectors acting on an object. This objective is achieved through adding (or subtracting) the vectors under consideration. Multiplication of vectors will not be presented here.

1.7.1 Properties of Vectors

There are several properties and guidelines that need to be considered when an operation on vectors is executed. Some of these guidelines are conventions that relate to how vectors are symbolically designated, and some are mathematical rules, which vectors must follow. The following is an outline of the most important rules:

1. The first relates to designating a vector in bold face in contrast to writing a scalar in regular script. For example, **A** refers to a vector that has a magnitude and a direction, while in general A is just a number that could be positive or negative. Some textbooks express a vector by a letter with an arrow over it. For example, $\vec{A}$ refers to the vector, **A**. The magnitude of a vector is indicated by placing it between two short vertical bars. For example, |**A**| indicates the magnitude of the vector, **A**. The magnitude of a vector is always a positive value.
2. Upon representing a vector, **A** (Figure 1.5a), graphically, an arrow in the form of a straight line, OL, with L representing the tip of the arrow, is drawn in the same direction as the vector (Figure 1.5b). Upon sketching the vector, an appropriate scale is chosen. In doing so, we are allowed to relocate the assigned vector to start from any position as long as its direction and magnitude are kept unchanged. The direction is indicated by the counter clockwise angle θ the vector makes with positive x axis. For example, the vector in Figure 1.5a was relocated for representation on a coordinate system (Figure 1.5b), while its magnitude and direction were kept fixed. An example is a vector that represents a displacement of 45 km in the horizontal direction, that is, east. This can be represented by the vector **A** (arrow OL) of 3.00 cm pointing in the x direction. The scale here is 1.0 cm (one block in Figure 1.6) for each 15 km. Another displacement of 50 km, directed 53° north of east, can be represented by a vector **B** (arrow OM) of 3.33 cm length that makes an angle of 53° with the positive x axis.
3. The negative of any vector, **B**, say, is a vector of a magnitude equal to that of **B**, but in a direction opposite to **B** (Figure 1.7). The negative of **B** in the above example, that is (–**B**), is a vector that makes an angle of 180° with the vector **B**, that is, an angle of 180° + 53° (= 233°), measured counterclockwise from the positive x axis.
4. If two vectors are equal, then they are equal in magnitude and direction. Therefore, if two vectors, of equal magnitude, happened to be parallel, then they are considered identical. Accordingly, vectors, **A**, **B**, and **C** (Figure 1.8), that represent some displacement, are all identical vectors. Regardless of the fact that they, as sketched, are in different locations, they represent the same displacement. As the head of any vector has coordinates that signify its x and y components in a Cartesian coordinate system with origin at the tail of the vector, then if two vectors are equal, their x components are equal and so are their y components.

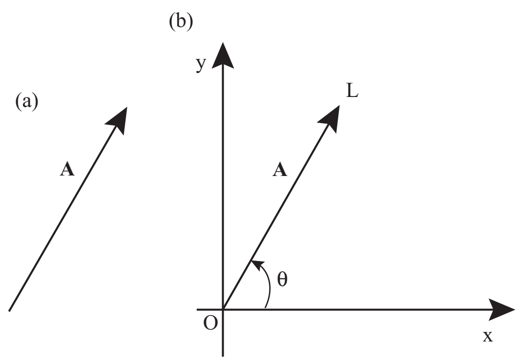


FIGURE 1.5 Demonstration of the property of relocating a vector, **A**, from its position in (a) to another position in (b) as long as the magnitude and direction of **A** are kept unchanged.

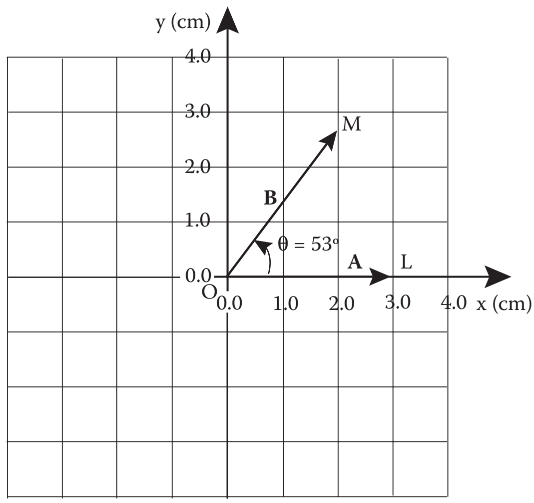


FIGURE 1.6 Representation of two displacements, 45 km east and 50 km 53° north of east.

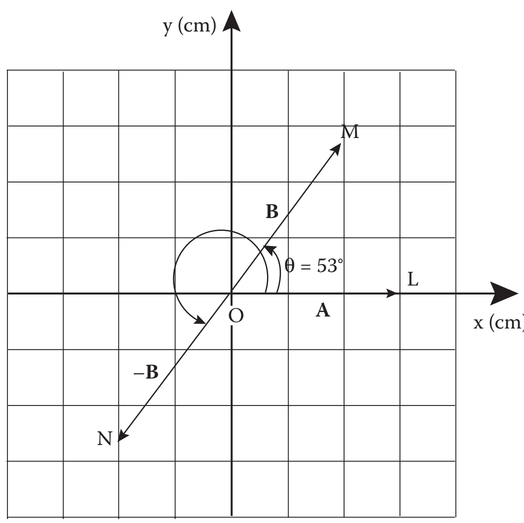


FIGURE 1.7 An example showing that the negative of a vector **B** is equal in magnitude but opposite in direction to the vector **B**.

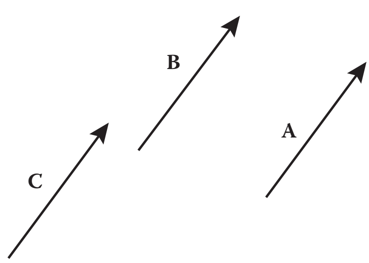


FIGURE 1.8 The three vectors **A**, **B**, and **C** are identical. All are parallel and have the same value.

1.8 ADDITION AND SUBTRACTION OF VECTORS

There are three techniques that can be followed in adding or subtracting vectors. These are

1. Graphical methods
2. Method by components
3. Analytical method

In this chapter, the first two methods are discussed.

1.8.1 Graphical Method

1.8.1.1 Addition

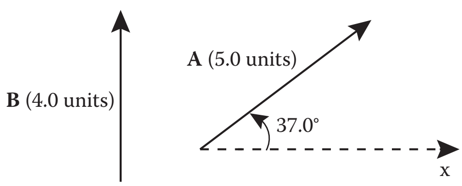
The addition of two vectors, **A** and **B**, is equivalent to finding the resultant, **R**, of these two vectors. Mathematically, this is represented by the vector equation

$$\mathbf{R} = \mathbf{A} + \mathbf{B}. \tag{1.4}$$

However, this addition can be demonstrated graphically for the two vectors, **A** and **B**, as stated in the following example.

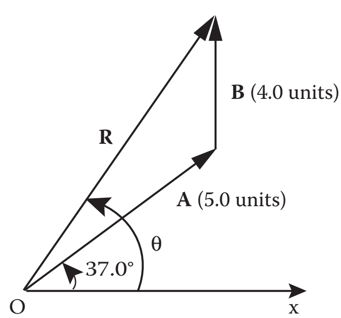
EXAMPLE 1.14

Determine graphically the resultant **R** of the two vectors **A** (5.0 units) and **B** (4.0 units), whose directions are described in the figure below.



Solution and Analysis

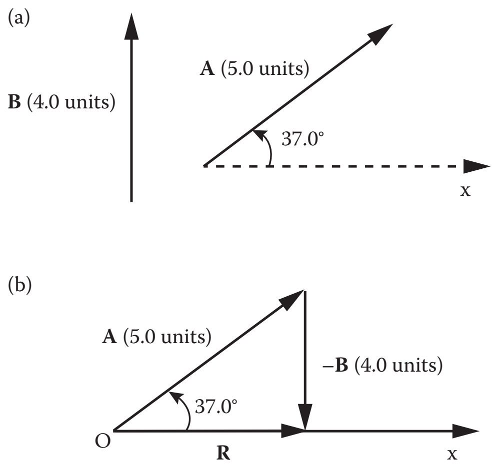
To determine the resultant **R** = **A** + **B** of the two vectors, one can take advantage of rule 2 (Section 1.7) that allows to relocate any vector as long as its magnitude and direction are kept fixed. So resketch **A** starting at another location, O (the figure below), and from the head of the arrow **A**, represent **B** by a line drawn parallel to **B**, that is, with an arrow pointing as **B** does, north. Connecting the starting point, O, with the head of the arrow, **B**, one obtains the resultant, **R**. So, in essence, after drawing the first vector, plotting the second vector, starting from where the first vector ended, and finally connecting the initial starting location, point O, with the last point (head of **B**), one obtains the resultant, **R**.



Measuring the length of the resultant **R** gives its magnitude, and measuring in a counterclockwise, the angle, which **R** makes with the x axis, counterclockwise, gives the direction of the resultant, **R**. During this process, the scale used in sketching the vectors in the figure above may be kept as that given in the figure under Example 1.14, that is, as that adopted for the given vectors **A** and **B** in the figure under Example 1.14 or choosing a new scale that either magnifies or demagnifies both vectors by the same ratio. If the scale in the figure above was kept the same as that of the figure under Example 1.14, one to one, then direct measurement of the length of **R** by a ruler yields directly the magnitude of the resultant vector **R**, and measurement of the angle it makes with the x axis yields the vector's direction.

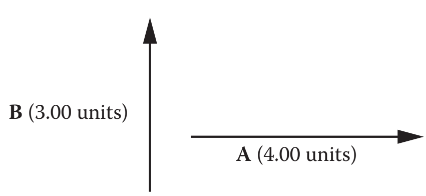
1.8.1.2 Subtraction

Subtraction of two vectors is treated in a way similar to that followed in the addition process, except noting that a negative vector is opposite in direction to that same vector, but both are equal in magnitude. So the operation **R** = **A** – **B** can be considered as **R** = **A** + (–**B**), which is basically adding the two vectors **A** and (–**B**). Following the same procedure, we can see that the two vectors and their subtraction can be represented in the diagram below, respectively.



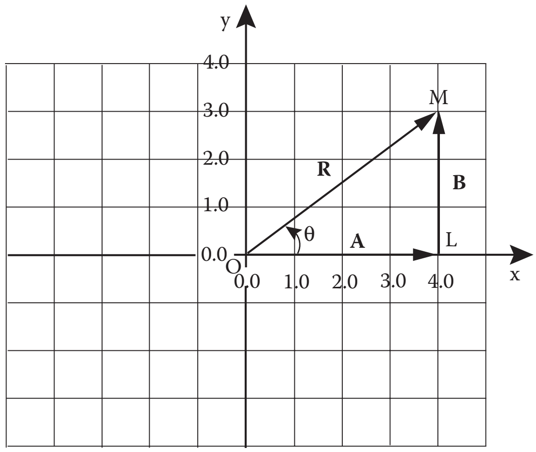
EXAMPLE 1.15

Determine graphically the resultant **R** of the two vectors **A** (4.0 units) directed east and **B** (3.0 units) directed north. The vectors directions are described in the figure below.



Solution and Analysis

Following the rules of addition, head-to-tail method, we start with OL representing the vector **A**, and from the head of vector **A** at L draw LM representing the vector **B**. The line OM then is the resultant **R** (the following figure). Direct measurement of the length of **R** by a ruler yields directly the magnitude of the resultant vector **R**, and measurement of the angle it makes with the x axis yields its direction.



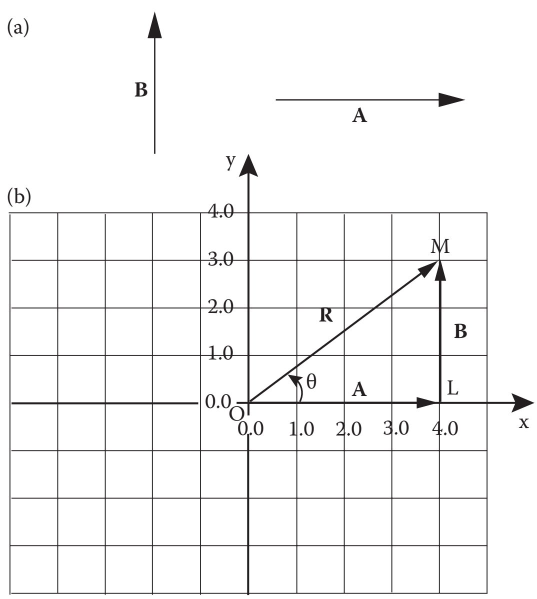
1.8.2 Addition by the Method of Components: General Treatment

In presenting this method, we begin with a special case of addition of the two vectors that was discussed in Example 1.15.

The connection between Example 1.15 and the method of components is served through the two vectors **A** and **B** intentionally directed along x and y, respectively. These, depicted in the figure below (part (a)), and added graphically, part (b), resulted in the construction of a right angle triangle whose hypotenuse is equal to the magnitude of the resultant **R**. The resultant along the hypotenuse is called **V**, just to reflect an arbitrarily chosen symbol for a vector **V**.

The triangle OLM in part (b) is a right angle triangle in which the length of the vector **A**, labeled in the figure as OL, is equal to A = OM cos θ, and the length of the vector, **B**, labeled in the figure as is equal to **B** = OM sin θ. That is,

$$\mathbf{A} = \mathrm{OM} \cos \theta, \quad \mathbf{B} = \mathrm{OM} \sin \theta.$$



Since OM is the length of the resultant, that is, OM = **V**, then the above relations can be rewritten as

$$A = V\cos\theta, \tag{1.5}$$

$$B = V\sin\theta. \tag{1.6}$$

Since **A** and **B** are along the x and y axes, respectively, then **A** is the x component of **V** and **B** is the y component of the **V**. In other words, the vector addition

$$\mathbf{V} = \mathbf{A} + \mathbf{B}$$

can equivalently be considered as the addition of the two component vectors, $\mathbf{V}_x$ and $\mathbf{V}_y$, in the vector equation

$$\mathbf{V} = \mathbf{V}_x + \mathbf{V}_y, \tag{1.7}$$

where, component wise,

$$V_x = V\cos\theta, \tag{1.8a}$$

$$V_y = V\sin\theta. \tag{1.8b}$$

That is, any general vector **V** can be expressed in a vector form as the addition of its Cartesian components. There are two components for a two-dimensional vector and three components for a three-dimensional vector.

It has to be noted that the Cartesian components of any vector are independent of each other, and hence V_x is independent of V_y and vice versa. The magnitude of the vector **V** is determined by the Pythagorean theorem, which gives

$$V = \sqrt{V_x^2 + V_y^2} \tag{1.9}$$

and the direction of the vector **V** is given by an angle θ obtained from dividing Equation 1.8b by Equation 1.8a. Thus,

$$\theta = \tan^{-1}\left(\frac{V_y}{V_x}\right). \tag{1.10}$$

Equations 1.8 through 1.10 hold for any vector, and hence knowing the magnitude and direction of the vector, one can use Equations 1.8 to resolve the vector into its components, and if the components of the vector are given, Equations 1.9 and 1.10 can be used to determine the magnitude of the vector and the angle it makes with the positive x-axis measured counterclockwise.

EXAMPLE 1.16

Determine the resultant of the addition of two vectors **A** (4.0 units) and **B** (3.0 units) using the method of components. **A** is along the x-axis and **B** is perpendicular to **A**.

Solution

Following the conclusions established at the end of the previous example, we calculate the x and y components of both **A** and **B** and relate the addition of the x components of **A** and **B** to the x component of the resultant. We repeat this process for the y components as well. For such a procedure, it would be helpful to construct the following table:

Vector	x Components	y Components
A (A = 4.0 units; θ = 0)	$A_x = A \cos\theta$	$A_y = A \sin\theta$
	$= 4.0 \cos 0°$	$= 4.0 \sin 0°$
	$= 4.0$	$= 0.0$
B (B = 3.0 units; θ = 90)	$B_x = B \cos\theta$	$B_y = B \sin\theta$
	$= 3.0 \cos 90°$	$= 3.0 \sin 90°$
	$= 0.0$	$= 3.0$
R (R = ?; θ = ?)	$R_x = A_x + B_x$	$R_y = A_y + B_y$
To be determined	$= 4.0 + 0.0$	$= 0.0 + 3.0$
	$= 4.0$	$= 3.0$

Thus, the magnitude of the resultant $\mathbf{R} = \sqrt{R_x^2 + R_y^2} = \sqrt{(4.0)^2 + (3.0)^2} = 5.0$ units, and direction of the resultant **R** is $\theta = \tan^{-1}(R_y/R_x) = \tan^{-1}(3.0/4.0) = 37°$.

EXAMPLE 1.17

Given the two vectors **A** (4.0 units) and **B** (3.0 units) in Example 1.16, determine the magnitude and direction of the vector **R** = **A** – **B**.

Solution

Taking advantage of the calculations made in the previous example, gives the following table:

Vector	x Components	y Components
A (A = 4.0 units; θ = 0)	$A_x = A \cos\theta$	$A_y = A \sin\theta$
	$= 4.0 \cos 0°$	$= 4.0 \sin 0°$
	$= 4.0$	$= 0.0$
B (B = 3.0 units; θ = 90)	$B_x = B \cos\theta$	$B_y = B \sin\theta$
	$= 3.0 \cos 90°$	$= 3.0 \sin 90°$
	$= 0.0$	$= 3.0$
R (R = ?; θ = ?)	$R_x = A_x$ ⊠ B_x	$R_y = A_y$ ⊠ B_y
To be determined	= 4.0 ⊠ 0.0	= 0.0 ⊠ 3.0
	$= 4.0$	= ⊠3.0

Thus, the magnitude of the resultant

$$\mathbf{R} = \sqrt{R_x^2 + R_y^2} = \sqrt{(4.0)^2 + (3.0)^2} = 5.0 \text{ units},$$

and the direction of the resultant **R** is

$$\theta = \tan^{-1}\left(\frac{R_y}{R_x}\right) = \tan^{-1}\left(\frac{-3.0}{4.0}\right) = -37°.$$

Analysis

1. As one can see, the magnitude of the resultant in this example is equal to the magnitude of the resultant in the previous example. However, the direction of the resultant here as given by a typical calculator is –37°. Therefore, the resultant in this example and the resultant in the previous example are two different vectors.
2. As the direction of the resultant in this example as given by a typical calculator is –37°, a further look is needed. This angle, taken on its value, implies that the resultant is a vector that lies in the fourth quadrant. Since the components of the resultant as calculated were $\mathbf{R}_x = 4.0$ and $\mathbf{R}_y = -3.0$, the resultant has to be in the fourth quadrant. Since we prefer to give the direction of a vector by the angle it makes with the positive x axis, the direction of the resultant being –37° is equivalent to (360° – 37°) = 323°. Thus, the resultant here makes an angle of 323° with the positive x axis.

1.9 GUIDELINE FOR DETERMINING THE DIRECTION OF THE RESULTANT

The direction of a two-dimensional vector **R** is determined from the relation

$$\theta = \tan^{-1}\left(\frac{R_y}{R_x}\right).$$

Upon substituting for the values of R_y and R_x, with their signs taken into consideration, one encounters one of the following cases:

1. The x component and the y component are both positive. In this case, the angle θ given on the calculator is a positive number and is the correct angle that describes the direction of the vector with respect to the positive x-axis (Figure 1.9a).
2. The x component is negative while the y component is positive. In this case, the angle as calculated on the calculator is a negative number, say (–45°). Since the components

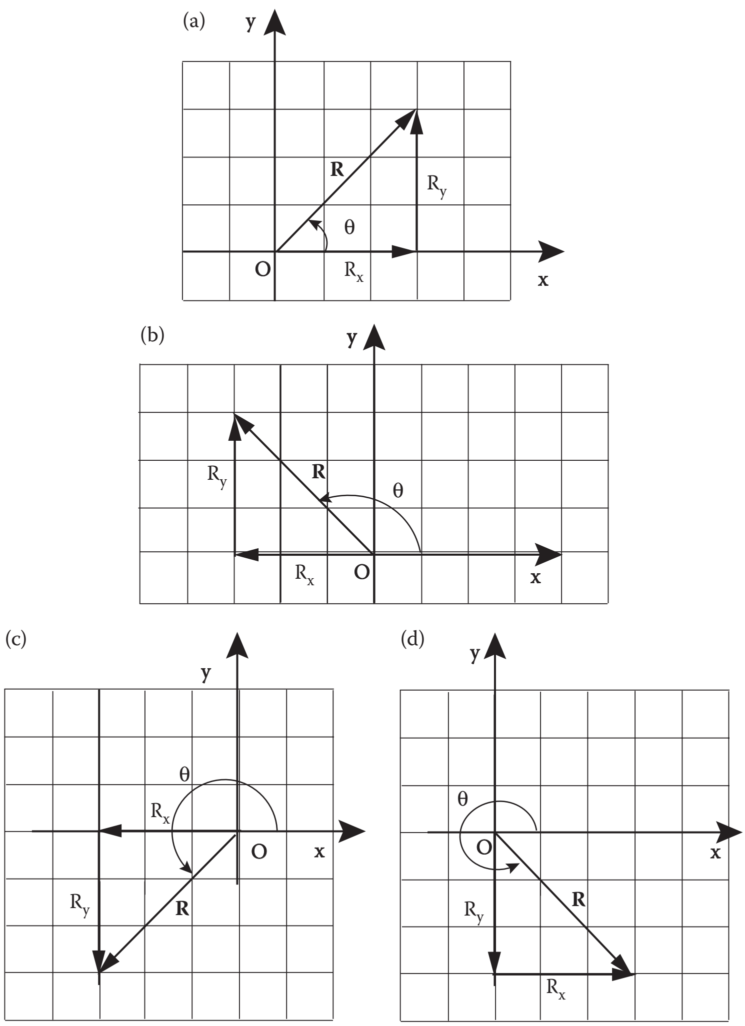


FIGURE 1.9 Four possible angles that a resultant vector could have according to the values and signs of its x and y components.

of this vector are x negative and y positive, the vector should lie in the second quadrant. Accordingly, the correct value of the angle should be 180° – 45°. That is, in this case, $\theta = 135°$ (Figure 1.9b).

3. The x component and the y component are both negative. In this case, the angle displayed on the calculator is a positive number, say 45°. Since the components, x and y, of this vector are both negative, the vector should lie in the third quadrant. Accordingly, the correct value of the angle should be 180° + 45°. That is, in this case, $\theta = 225°$ (Figure 1.9c).
4. The x component is positive while the y component is negative. In this case, the angle as displayed on the calculator is a negative number, say (–45°). Since the components of this vector are: x positive, and y negative, the vector should lie in the fourth quadrant. Accordingly, the correct value of the angle should be 360° – 45°. That is in this case, $\theta = 315°$ (Figure 1.9d).

EXAMPLE 1.18

Repeat Example 1.14 and determine the resultant of the two given vectors **A** (5.00 units) that makes an angle of 37° with the positive x axis and **B** (4.00 units) that points to north, that is, it makes an angle of 90° with the positive x axis (see figure under Example 1.14).

Solution

Following the procedure used in the previous example, consider the following table:

Vector	x Components	y Components
A (A = 5.0 units; θ = 37)	$A_x = A \cos\theta$ $= 5.00 \cos 37°$ $= 4.00$	$A_y = A \sin\theta$ $= 5.00 \sin 37°$ $= 3.00$
B (B = 3.0 units; θ = 90)	$B_x = B \cos\theta$ $= 4.00 \cos 90°$ $= 0.00$	$B_y = B \sin\theta$ $= 4.00 \sin 90°$ $= 4.00$
R (R; θ) To be determined	$R_x = A_x + B_x$ $= 4.00 + 0.00$ $= 4.00$	$R_y = A_y + B_y$ $= 3.00 + 4.00$ $= 7.00$

Thus, the magnitude of the resultant

$$\mathbf{R} = \sqrt{R_x^2 + R_y^2} = \sqrt{(4.00)^2 + (7.00)^2} = 8.06 \text{ units},$$

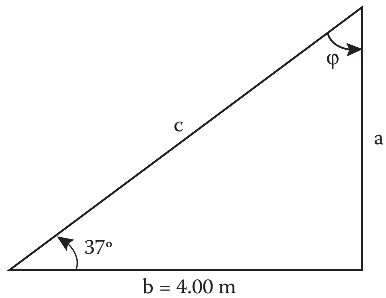
and the direction of the resultant **R** is

$$\theta = \tan^{-1}\left(\frac{R_y}{R_x}\right) = \tan^{-1}\left(\frac{7.00}{4.00}\right) = 60.3°.$$

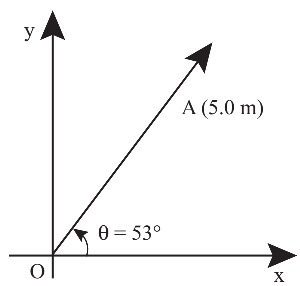
PROBLEMS

1.1 Carry out the following operations rounding off the final answer to the proper number of significant figures:
   a. The product: 0.80 × (4430)
   b. The division: 131/100
   c. The subtraction: 1.24 – 2.7

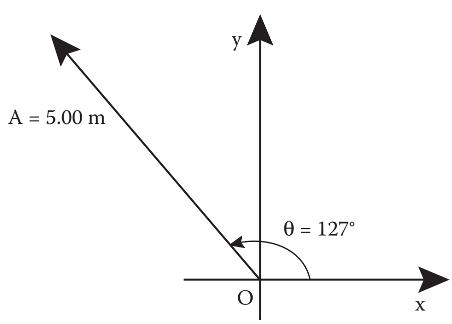
1.2 Write the following numbers in scientific notation and express each in four significant figures:
   a. 514,200
   b. 0.002460
1.3 Write the following numbers in scientific notation and express each in three significant figures:
   a. The mean radius of the Earth R is 6370.949 km*
   b. Bohr's radius for the hydrogen atom r is 0.0000000000529177249 m†
1.4 In the figure below, use trigonometry to determine (a) the angle φ, (b) the side a, and (c) the hypotenuse c.



1.5 Refer to the figure above and use the Pythagorean theorem to find the hypotenuse c. Does your answer verify your previous answer for part (c) in Problem 1.4?
1.6 Given the vector **A** in the figure below, determine its x and y components.

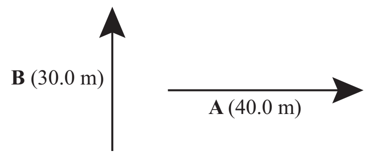


1.7 Given the vector **A** in the figure below, determine its x and y components.



* NASA Technical TRanslation NASA TT-F-533.
† B. N. Taylor, W. H. Paker, and D. N. Langenberg, *Rv. Mod. Phys.*, 41, 375, 1969.

1.8 Consider the two vectors depicted in the figure below. Use graphical method to determine the resultant $\mathbf{R} = \mathbf{A} + \mathbf{B}$.



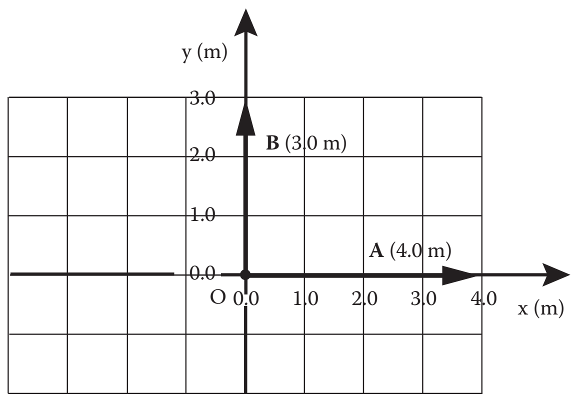
1.9 For the two vectors described in Problem 1.8, determine graphically the vector $\mathbf{R} = -(\mathbf{A} + \mathbf{B})$.

1.10 Given the two vectors the figure below, determine

a. The x and y components of each vector.

b. The magnitude and direction of the resultant vector $\mathbf{R}_1 = \mathbf{A} + \mathbf{B}$.

c. The magnitude and direction of the resultant vector $\mathbf{R}_2 = \mathbf{A} - \mathbf{B}$.



1.11 Refer to the figure above to determine the following vectors:

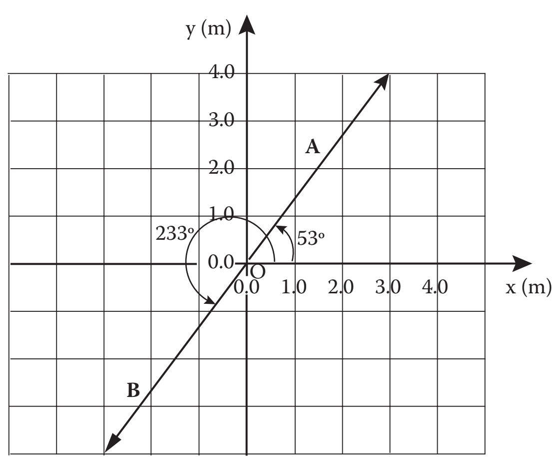
a. $\mathbf{R}_1 = -\mathbf{A} - \mathbf{B}$

b. $\mathbf{R}_2 = -\mathbf{A} + \mathbf{B}$

1.12 A vector **A** is of 5.00 units and makes an angle of 53.0° with the positive x-axis. Another vector **B** is of 5.00 units and makes an angle of 233° with the x-axis (the figure below). Using the method of components, determine the magnitude of the vectors $\mathbf{R}_1$ and $\mathbf{R}_2$ given by

a. $\mathbf{R}_1 = \mathbf{A} - \mathbf{B}$

b. $\mathbf{R}_2 = \mathbf{A} + \mathbf{B}$

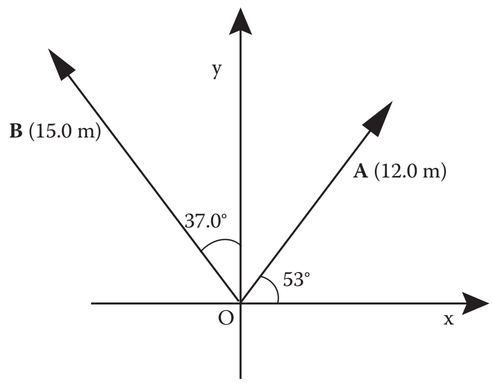


1.13 A vector **A** is of 4.00 units and makes an angle of 90° with the positive x-axis. Another vector **B** is of 4.00 units and makes an angle of 180° with the x-axis.

a. Determine the x and y components of each vector.
b. Determine the x and y components of the resultant **R** = **A** + **B**.
c. Using the method of components, determine the magnitude and direction of the resultant **R** = **A** + **B**.

1.14 Refer to the figure below to compute the following:

a. The x and y components of the vectors **A** and **B**
b. The components of the resultant **R** given by **R** = **A** + **B**
c. The magnitude of the resultant **R**
d. The direction of the resultant **R**



1.15 A boy scout in search of his friend walked 3.00 km east and followed this with a walk of 4.00 km south.

a. Use the method of components to determine the magnitude and direction of the boy's resultant displacement from the start. Call this resultant $\mathbf{R}_1$.
b. Using part (a) and without further calculations, determine the magnitude and direction of a short-cut displacement that would take the boy from his new position back to his starting position. Call this short-cut displacement $\mathbf{R}_2$.
c. Using parts (a) and (b) and with no further calculations, determine the magnitude and direction of the resultant $\mathbf{R} = \mathbf{R}_1 + \mathbf{R}_2$.

2 Motion in One Dimension

This chapter addresses deriving and using the equations of motion that describe the time dependence of an object's displacement, velocity, and acceleration. Relationships between displacement, velocity, and acceleration are also of importance and will be derived. This chapter starts with the basic definitions of displacement and average velocity of an object moving in one dimension. Such a simplified start will help to lead a complete set of equations of motion for an object moving along one dimension, east–west, north–south, or up and down. In the context of coordinate systems that were treated in the previous chapter, the one-dimensional motion will reduce the time and effort needed on the study of motion in two dimensions, which is the subject of the next chapter.

2.1 DISPLACEMENT

Motion can be defined as a continuous change in position and that change could occur in one, two, or three dimensions. As the treatment here will be limited to motions only in one dimension, one axis of the coordinate system described in Chapter 1 will suffice. Choosing this axis as the x-axis, we depict on it a point, O, which will be considered as an origin of zero coordinate. Any point on the right of the origin O will have a positive coordinate and any point on the left of O will have a negative coordinate (Figure 2.1).

The displacement, denoted by Δx, of an object as it moves from an initial position x_i to a final position x_f along the x-axis can be defined as the change in the object's position along the x-axis. That is

$$\Delta x = x_f - x_i. \tag{2.1}$$

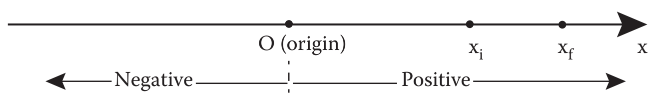


FIGURE 2.1 Depiction of position along the x-axis. O defines the origin where positive coordinates are on its right and negative coordinates are on its left.

2.1.1 Special Remarks

The quantities in Equation 2.1 are treated as vectors. x_i is the initial position vector, x_f the final position vector, and Δx the displacement vector. Accordingly, all these quantities have a direction and magnitude. Since they are all one-dimensional, the sign of each quantity is a sufficient description of the quantity pointing along the positive or negative directions of the x-axis.

EXAMPLE 2.1

Determine the displacement of an object that has moved from an initial position x_1 to a final position x_2 given by the following cases:

a. $x_1 = 4.0$ m, $x_2 = 9.0$ m
b. $x_1 = 4.0$ m, $x_2 = -9.0$ m
c. $x_1 = -4.0$ m, $x_2 = -9.0$ m

Solution

Following the definition as given by Equation 2.1, we then have

a. $\Delta x = x_2 - x_1 = 9.0\text{ m} - 4.0\text{ m} = 5.0\text{ m}$
b. $\Delta x = x_2 - x_1 = -9.0\text{ m} - 4.0\text{ m} = -13.0\text{ m}$
c. $\Delta x = x_2 - x_1 = -9.0\text{ m} - (-4.0\text{ m}) = -9.0 + 4.0 = -5.0\text{ m}$

Analysis

As one can observe from the above values, the displacement may be positive or negative depending on the object's initial and final positions. The positive sign simply means that the net movement of the object occurred along the positive x-axis or eastward, whereas the negative sign simply means that the net movement occurred along the negative x-axis or westward.

EXAMPLE 2.2

Determine the net displacement of an object that has executed two successive displacements, one from position $x = 1.0$ m to position $x = 7.0$ m, and another from this position back to its original position where $x = 1.0$ m.

Solution

Following the definition as given by Equation 2.1, we can solve this problem in one of the following two ways:

1. *First method:* First, we determine the displacement in each motion and add it algebraically to get the net displacement.
   The displacement in the forward part of the motion is

$$\Delta x_{forward} = x_f - x_i = 7.0\text{ m} - 1.0\text{ m} = 6.0\text{ m}.$$

   The displacement in the backward part of the motion is

$$\Delta x_{backward} = x_f - x_i = 1.0\text{ m} - 7.0\text{ m} = -6.0\text{ m}.$$

   The total displacement is the algebraic addition of the two displacements. Thus

$$\Delta x_{total} = \Delta x_{forward} + \Delta x_{backward} = 6.0\text{ m} + (-6.0\text{ m}) = 0\text{ m}.$$

2. *Second method:* We can look at the two motions, forward and backward, as the whole motion for which the initial position was $x_i = 1.0$ m and the very final position was also 1.0 m. As the total displacement in this motion is

$$\Delta x_{total} = x_f - x_i,$$

   then substituting for the given initial and very final position values, we get

$$\Delta x_{total} = 1.0\text{ m} - 1.0\text{ m} = 0.0\text{ m}.$$

Analysis

1. What matters in calculating the total displacement is the very initial and very final position values. Any intermediate position stops do not influence the final answer.
2. The net displacement in this motion being zero does not mean that the distance traveled from the initial value to the very final value is zero. In fact, it is 6.0 m for each part of the motion. Accordingly, the total distance, d, the object has traveled is given by

$$d = 6.0\text{ m} + 6.0\text{ m} = 12.0\text{ m}.$$

The above example presents to us the first case in which we see an important difference between displacement and distance. Displacement, in essence, is a vector quantity that could take positive, negative, or zero values. But distance is always a positive number.

2.2 AVERAGE VELOCITY

The average velocity of an object moving between two positions in a certain period of time is an abstract quantity that gives an idea of what constant velocity it would need to travel through the distance separating the two positions in the same period of time. This is an ideal notion, since most motions are not constant in time. However, the concept of the average velocity as will be seen shortly is very useful.

The average velocity of an object moving between two locations in the time interval, Δt, is defined as

$$\bar{v} = \frac{\Delta x}{\Delta t}; \quad \Delta x = x_f - x_i; \quad \Delta t = t_f - t_i \tag{2.2}$$

or

$$\Delta x = \bar{v}\Delta t \tag{2.3}$$

or, assuming that the start is set at $t_i = 0$, then

$$\Delta x = \bar{v}t. \tag{2.4}$$

This is graphically projected in Figure 2.2 in which point P designates the initial location and point Q designates the final destination. Notice that the curve between points P and Q represents the actual path of the object's motion. However, the dotted straight line connecting points P and Q is a short-cut path that alternatively tells one if the object was going along this short cut, it would reach its destination Q in the same time interval Δt that it actually took the object to get to Q. In the x–t plot, $\Delta x/\Delta t$ is simply the slope (rise/run) of the dotted, straight line PQ.

It should be noted that since $\mathbf{\Delta x}$ in Equation 2.2 is a vector, the average velocity is a vector and connecting the possible values that $\mathbf{\Delta x}$ could take, the average velocity of an object during a certain motion could assume positive, negative, or zero values. It is the sign of the displacement $\mathbf{\Delta x}$ that dictates the sign of the average velocity.

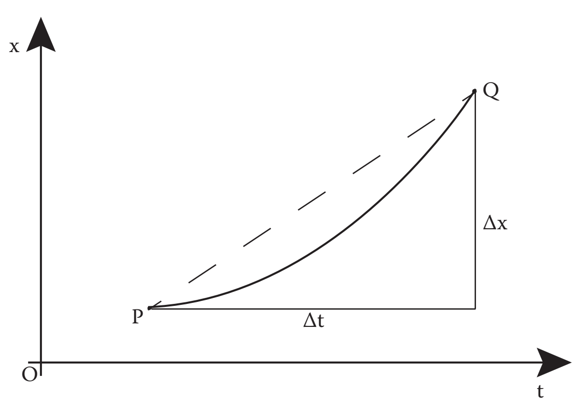


FIGURE 2.2 An x–t plot for an object between its initial location at P and final destination at Q. Δx is called the "rise" of the line PQ, whereas Δt is its "run." The average velocity is simply $\bar{v}$ = rise/run = $\Delta x/\Delta t$.

2.3 AVERAGE SPEED

Along with the notion of velocity, there is another notion, speed, which is in common use. This is always a positive number that has no relevance to the direction in which the object is moving. The speed of an object traveling a distance, Δd, in a time interval Δt is defined as

$$v = \frac{\Delta d}{\Delta t}.$$

If the distance $\Delta d = d_f - d_i$ is measured from a zero-position reference ($d_i = 0$) and Δt is measured with an initial time setting chosen as zero ($t_i = 0$), then $\Delta d = d_f = d$ and $\Delta t = t_f = t$. In this case, the above equation becomes

$$v = \frac{d}{t}. \tag{2.5}$$

In the case of motor vehicles, the speed is what is read on the speedometer of the car and it is the value for which a driver could be ticketed if he was driving over the speed limit.

EXAMPLE 2.3

Consider an object that has executed two successive displacements, one from position $x = 1.0$ m to position $x = 7.0$ m in a time interval of 5.0 s, and another from this position back to its original position, $x = 1.0$ m, in 10.0 s. Determine the average velocity

a. In the forward part of the motion
b. In the backward part of the motion
c. In the whole trip

SOLUTION

Following the definition as given by Equation 2.2, we see that we need to calculate the displacement for each case of the motion.

a. The displacement in the forward part of the motion is

$$\Delta x_{forward} = x_f - x_i = x_2 - x_1 = 7.0 \text{ m} - 1.0 \text{ m} = 6.0 \text{ m}.$$

As the time interval for this part of the motion was 5.0 s, the average velocity of the object in the forward part of the motion is

$$\bar{v} = \frac{\Delta x}{\Delta t} = \frac{6.0 \text{ m}}{5.0 \text{ s}} = 1.2 \text{ m/s}.$$

b. The displacement in the backward part of the motion is

$$\Delta x_{backward} = x_f - x_i = x_2 - x_1 = 1.0 \text{ m} - 7.0 \text{ m} = -6.0 \text{ m}.$$

As the time interval for this part of the motion was 10 s, then the average velocity of the object in the backward part of the motion is

$$\bar{v} = \frac{\Delta x}{\Delta t} = \frac{-6.0 \text{ m}}{10.0 \text{ s}} = -0.60 \text{ m/s}.$$

c. The total displacement of the two successive displacements is

$$\Delta x_{total} = \Delta x_{forward} + \Delta x_{backward} = 6.0 \text{ m} + (-6.0 \text{ m}) = 0.0 \text{ m}.$$

As the time interval for the whole round trip motion was 15.0 s, then the average velocity of the object in the round trip motion is

$$\bar{v} = \frac{\Delta x}{\Delta t} = \frac{0.0\ \text{m}}{15.0\ \text{s}} = 0.0\ \text{m/s}.$$

Analysis

1. Notice that the average velocity was positive in part (a), was negative in part (b), and zero in part (c). This asserts our discussion of the displacement and how its direction dictates the direction of the average velocity.
2. From the values given in this example, one observes that the total distance covered in this whole motion is equal to

$$d = 12.0\ \text{m}.$$

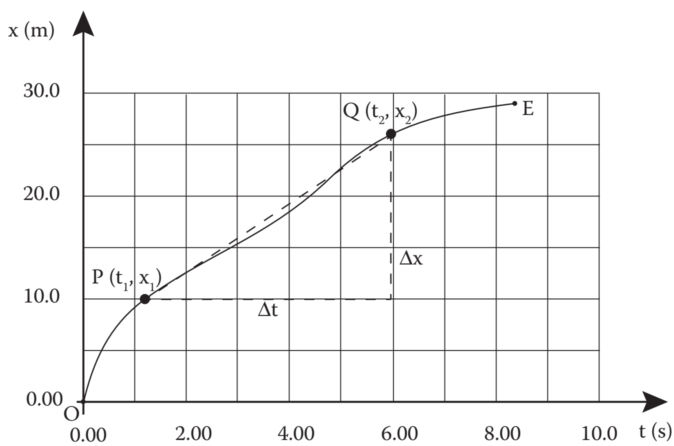
Therefore, the average speed (always a positive value) is

$$v = \frac{d}{t} = \frac{12.0\ \text{m}}{15.0\ \text{s}} = 0.800\ \text{m/s}.$$

One may establish a rather simple connection between the notion of speed and the reading on the speedometer of a car. The reading is always a number, whereas velocity in addition to its value carries the sense of direction.

EXAMPLE 2.4

In this example, a traffic motorist, starting from a location represented in the diagram by the origin, O, headed on a straight highway toward his watching postdestination at location E. The x–t plot of the motorist is depicted in the figure below. Determine the average velocity of the motorist during the PQ part of the motion.



Solution

Connecting the points P and Q, and constructing a right angle triangle by dropping a vertical line from Q and drawing a horizontal line from P, gives the triangle with sides Δx and Δt and gives PQ the hypotenuse.

From the diagram, the displacement is

$$\Delta x = x_2 - x_1 = 26\ \text{m} - 10\ \text{m} = 16\ \text{m},$$

and the time interval is

$$\Delta t = t_2 - t_1 = 5.9\ \text{s} - 1.2\ \text{s} = 4.7\ \text{s}.$$

As the average velocity is given by

$$\bar{v} = \frac{\Delta x}{\Delta t},$$

then for this motorist

$$\bar{v} = \frac{\Delta x}{\Delta t} = \frac{x_2 - x_1}{t_2 - t_1} = \frac{16\ \text{m}}{4.7\ \text{s}} = 3.4\ \text{m/s}.$$

2.4 INSTANTANEOUS VELOCITY

This is another velocity concept that describes the velocity magnitude and direction of an object at a certain instant of time during the object's motion. In Figure 2.3a, the x–t diagram describes an object's motion between two points P and Q, with numerous positions, P_1, P_2, P_3, and so on in-between. The instantaneous velocity of the object at a certain instant, say P_1, can be introduced through the concept of average velocity calculated in an infinitesimal segment of the motion such as P_1P_2 and in a manner similar to that introduced in Figure 2.2. A right angle triangle is constructed such that the chord P_1P_2 becomes the hypotenuse. The sides Δx and Δt that represent the displacement and the time interval for this segment of the motion are infinitesimally small and the line connecting P_1–P_2 becomes tangent to the curve at P_1 (Figure 2.3b). The slope of this tangent is

$$v = \lim_{\Delta t \to 0} \left(\frac{\Delta x}{\Delta t} \right). \tag{2.6}$$

Equation 2.6 is the mathematical definition of the instantaneous velocity at point P_1.

Accordingly, to calculate the instantaneous velocity at some instant of an object's motion when an x–t diagram is given, one only needs to determine the slope of the line tangent to the curve at the point of interest. For example, in the x–t plot in Figure 2.4, the instantaneous velocity at point M is

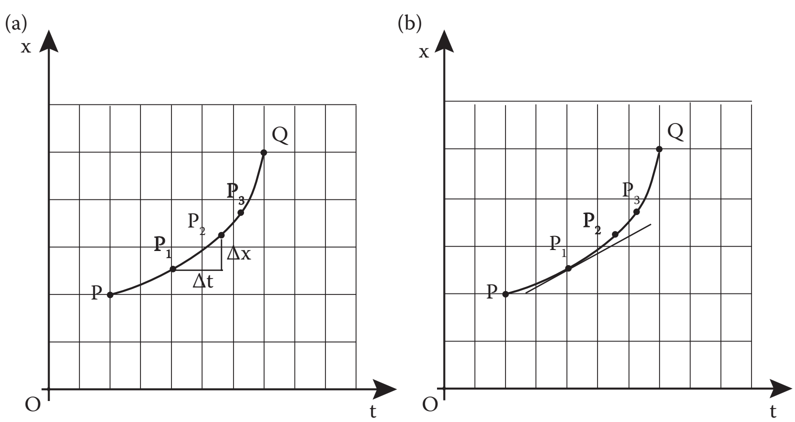


FIGURE 2.3 (a) An x–t plot that describes an object's motion between two positions P and Q; P_1 P_2, P_2 P_3, and so on are infinitesimally small segments along the object's path; (b) the instantaneous velocity of the object at P_1 is the slope of the line tangent to the plot at P_1.

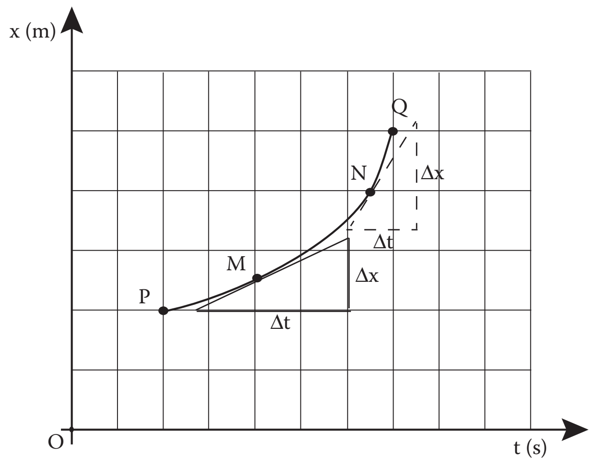


FIGURE 2.4 The tangent at two points M and N of an object's x–t plot. In each case, a right angle triangle is constructed such that the hypotenuse is tangent to the point at which the slope of the tangent gives the instantaneous velocity of the object.

equal to the slope of the line tangent to the plot at M. The slope is usually denoted by m and from the well-known expression, this is equal to the (rise/run) of that line.

For the tangent line at M, the rise is Δx and the run is Δt. Notice that the rise and run are the two sides of the solid line in the right angle triangle drawn such that the tangent line at M is the hypotenuse. If one is interested in determining the instantaneous velocity at another instant of the motion represented by point N on the plot, then the instantaneous velocity of the object at N is equal to the slope of the tangent of the curve at N (the dotted tangent). The slope of this line is given by a similar expression, slope = rise/run of the tangent line at N. Notice that the rise and run of the tangent at N are the two dotted sides on the triangle drawn; the tangent line at N is the hypotenuse.

It should be noted that since Δx in Equation 2.6 is a vector, the instantaneous velocity is also a vector and considering the possible values that Δx could take, the instantaneous velocity of an object at any instant during motion could assume positive, negative, or zero values. It is the slope of the tangent that dictates the sign and value of the instantaneous velocity.

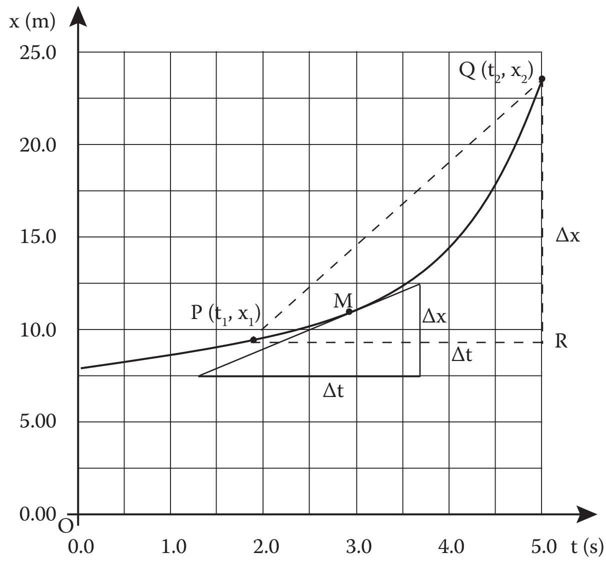
2.4.1 Special Remarks

The sign of the instantaneous velocity is an indication in what direction relative to the positive x-axis the object is moving at the instant of interest. Positive sign implies that the object is moving along the positive x-axis away from the origin. However, if the instantaneous velocity is negative, the object is moving in the negative x-direction, but its position could be either (+x) or (–x). In cases when the instantaneous velocity at a certain instant is zero, the object at that instant is stationary.

EXAMPLE 2.5

The following figure represents an x–t plot of a runner headed toward her destination at Q. The points P and M refer to two selected positions passed by the cyclist 1.9 and 2.9 s after her start. Using your best estimate, you can read from the plot to determine

a. The average velocity of the runner between P and Q
b. The instantaneous velocity of the runner at point M, 2.0 s after passing location P



Solution

a. As discussed earlier (Section 2.2), we connect the points P and Q and construct a right angle triangle by dropping a vertical line from Q and drawing a horizontal line from P. These are the dashed lines QR and PR, respectively, and the hypotenuse is the dashed line PQ. From the dashed sides Δx and Δt, the average velocity in the interval between P and Q is

$$\Delta x = x_Q - x_P = 23.8 \text{ m} - 9.0 \text{ m} = 14.8 \text{ m},$$

and the time interval between P and Q is

$$\Delta t = t_Q - t_P = 5.0 \text{ s} - 1.9 \text{ s} = 3.1 \text{ s}.$$

As the average velocity is given by

$$v_{av} = \frac{\Delta x}{\Delta t},$$

then for this runner

$$v_{av} = \frac{\Delta x}{\Delta t} = \frac{x_2 - x_1}{t_2 - t_1} = \frac{14.8 \text{ m}}{3.1 \text{ s}} = 4.8 \text{ m/s}.$$

b. As also discussed earlier (Section 2.4) for calculating the instantaneous velocity, we draw a line tangent to the plot at the point of interest, M, in this case, and construct on this line a right angle triangle (solid lines) such that the tangent becomes its hypotenuse. The vertical side of this triangle corresponds to a position change Δx, and the horizontal side corresponds to time elapsed Δt. The slope, m, of this tangent gives the instantaneous velocity, v, of the runner at the instant $t = 2.9$ s. Of course, Δx and Δt should be read from the plot and since these readings are estimates, they are not going to be exact. But, to a large degree, they can be estimated fairly accurately. Thus

$$v_{av} = \frac{\Delta x}{\Delta t} = \frac{(12.5 - 7.5) \text{ m}}{(3.7 - 1.3) \text{ s}} = \frac{5 \text{ m}}{2.4 \text{ s}} = 2.1 \text{m/s}$$

Analysis

1. The value of the instantaneous velocity is different from the average velocity that the runner had during this motion, which is generally expected.
2. The sign of the instantaneous velocity is an indication in what direction relative to the *x*-axis the runner was moving at the instant $t = 2.9$ s. The positive sign implies that the runner was moving away from the origin along the positive direction of the x-axis.

2.5 ACCELERATION

The acceleration of a moving object in a given time interval results from a change in its velocity during the motion. The notion "average" and "instantaneous" applies to acceleration in a manner similar to that applied to velocity. However, in this chapter, the acceleration of objects will be considered constant in value. Therefore, the complications of determining the instantaneous acceleration of an object during its motion are avoided. The average acceleration $\bar{a}$ is considered equal to the constant instantaneous acceleration a. Thus, the instantaneous acceleration of an object that has moved through a displacement (Δx) in a period of time (Δt) can be calculated from the change in its velocity divided by the time interval during which the change in velocity takes place. Thus,

$$a = \frac{\Delta v}{\Delta t}. \tag{2.7}$$

That is,

$$a = \frac{v - v_o}{t - t_o}$$

where v_o is the object's initial velocity and v is its final velocity. Setting $t_o = 0$ and t for t_f, the above equation becomes

$$a = \frac{v - v_o}{t}. \tag{2.8}$$

As for the initial time setting, it is always preferable to start the time count from zero, and hence, the final time setting is the actual period of time that elapsed during this motion.

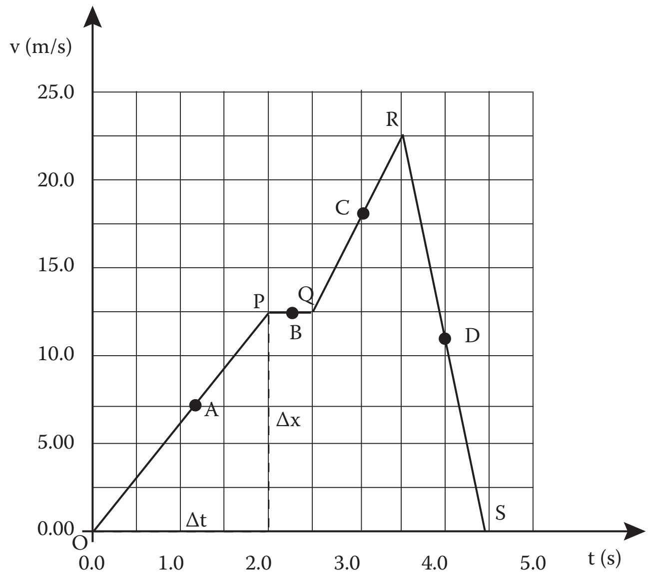
The above equation can be written as

$$v = v_o + at. \tag{2.9}$$

This is an equation of motion that enables one to determine the final velocity of an object at any instant of time t after its start with an initial velocity v_o and moving under an acceleration a.

EXAMPLE 2.6

Consider an object moving along the x direction and having the v–t plot depicted in the following figure. Determine the object's instantaneous acceleration at the instants labeled on the plot by points A, B, C, and D.



SOLUTION

As each of the segments OP, PQ, QR, and RS of the object's motion is a straight line, the slope of each line represents the average acceleration of the object in that segment. And as a straight line has the same slope at all of its points, the slope at A, B, C, and D equals the instantaneous acceleration at each of these points.

From Equation 2.8, for any segment of the object's motion, we have

$$a = \frac{v - v_o}{t}.$$

Thus, for each segment of the motion, considering v the final velocity of the object and v_o its initial velocity at the start of the segment, the instantaneous accelerations are

$$\text{At A:} \quad a = \frac{\text{rise}}{\text{run}} = \frac{(12.5 - 0.0)\ \text{m/s}}{(2.0 - 0.0)\ \text{s}} = 6.3\ \text{m/s}^2,$$

$$\text{At B:} \quad a = \frac{\text{rise}}{\text{run}} = \frac{(12.5 - 12.5)\ \text{m/s}}{(2.5 - 2.0)\ \text{s}} = 0.0\ \text{m/s}^2,$$

$$\text{At C:} \quad a = \frac{\text{rise}}{\text{run}} = \frac{(22.5 - 12.5)\ \text{m/s}}{(3.5 - 2.5)\ \text{s}} = 10.0\ \text{m/s}^2,$$

$$\text{At D:} \quad a = \frac{\text{rise}}{\text{run}} = \frac{(0.0 - 22.5)\ \text{m/s}}{(4.5 - 3.5)\ \text{s}} = -22.5\ \text{m/s}^2.$$

ANALYSIS

As noticed, since the sign of a at A was positive, the object's velocity was increasing in value along the positive x-axis. However, at B, the acceleration is zero, which means that the object kept moving toward Q with the same velocity it had at P. In the third segment QR, the sign of *a* at C was positive. Then the object's velocity was increasing in value along the positive x-axis, again getting farther away from the origin at a higher rate than that in the previous two segments. Finally, at D, the sign of *a* was negative and that means the object's velocity was decreasing in value along the positive x-axis and kept diminishing to zero when the object came to a final stop at t = 4.5 s.

2.6 EQUATIONS OF MOTION

Equation 2.9 is the first equation of motion that involves instantaneous velocities, initial, and final, and Equation 2.3 is another equation that describes the displacement in terms of the average velocity. Since acceleration, defined as the change in velocity with respect to time, is going to be considered always constant, the value of average velocity lies exactly in the middle between the values of the initial and final velocities. That is, the average velocity during a motion that is subject to a constant acceleration is just the average of the initial and final velocity. Thus,

$$\bar{v} = \frac{1}{2}(v_o + v). \tag{2.10}$$

Using Equation 2.3, the displacement is

$$\Delta x = \frac{1}{2}(v_o + v)t.$$

Substituting for v from Equation 2.9 into the above equation yields

$$\Delta x = \frac{1}{2}(v_o + (v_o + at))t,$$

which after some algebra becomes

$$\Delta x = v_o t + \frac{1}{2}at^2. \tag{2.11}$$

Equation 2.11 is the second equation of motion.

2.6.1 Special Remarks

1. It has to be noted that (Δx) in Equation 2.11 is a vector quantity and since this chapter addresses one-dimensional motion, then in solving for Δx, the vector aspect of Δx shows up in the sign its value takes. It could have a positive or negative sign, or even zero, depending on the case under consideration.
2. In a simplified way of tackling one-dimensional problems, it is so practical to lay out the setup of the problem such that the motion starts at $x_o = 0$. In this case, $\Delta x = x - 0 = x$, and Equation 2.11 becomes

$$x = v_o t + \frac{1}{2}at^2. \tag{2.12}$$

Again, for an object whose motion starts at $x_o = 0$, x would be the displacement and not the distance that the object has traveled in the interval t.

Equations 2.9 and 2.12 are the two important equations that are very helpful in describing an object's position and final velocity if the object's initial conditions and its acceleration are known.

Time is an explicit variable in Equations 2.9 and 2.12. A new equation can be generated from these two equations with no explicit time dependence on it. The time t can be eliminated by substituting for it from Equation 2.9 into Equation 2.12, giving the third equation of motion

$$v^2 = v_o^2 + 2\,ax. \tag{2.13}$$

Equations 2.3, 2.9, 2.12, and 2.13 offer a complete set of equations that can provide a detailed history of the object's motion once its initial conditions and acceleration are given.

EXAMPLE 2.7

An automobile is moving with a velocity of 20.0 m/s (45.0 mi/h) when its driver suddenly noticed an accident blocking a major part of the highway. He immediately pressed on the brakes and was able to bring his car to a full stop in a distance of 30.0 m. Determine

a. The acceleration of the car during this controlled slowing process
b. The time that it took the driver to get his car fully stopped from the moment he slammed on the brakes

Solution

a. Equation 2.13 is used here since it does not contain time, which remains unknown, but it does contain the acceleration.

$$v^2 = v_o^2 + 2\,ax.$$

Substituting for $v = 0$, $v_o = 20.0$ m/s, and $x = 30.0$ m, gives

$$0 = (20.0\text{ m/s})^2 + 2\,a\,(30.0\text{ m}).$$

Solving for a gives

$$a = -(4.00 \times 10^2)/60.0 = -6.67\text{ m/s}^2.$$

b. Now, the acceleration is known; so, to find the time taken to stop the car, use Equation 2.9.

$$v = v_o + at.$$

Substituting for the given quantities gives

$$0 = 20.0\text{ m/s} + (-6.67\text{ m/s}^2)t.$$

Solving for t yields

$$t = 3.00\text{ s}.$$

Analysis

1. The acceleration turned out to be negative. This means that the acceleration is in fact decelerating the car instead of accelerating it. That is, the change happening to the velocity of the car is from higher to lower velocities, which is what is expected in a case like this.
2. A value of 20.0 m/s (45.0 mi/h) for the initial velocity is considered a reasonable value. That is why the driver was able to stop within a distance of 30.0 m (98.4 ft) and in a time interval of 3.0 s.

2.6.2 General Remarks

1. As the treatment presented so far has been for one-dimensional motion, a coordinate basis for this motion, chosen in all of the previous equations to be the x-axis, should work if either the y- or the z-axis was to be adopted. Of course, all vector quantities in those equations would then refer to either the y- or the z-axis, respectively. It is just a matter of

convenience to adopt the axis that relates to the object's direction of motion. For example, in the free fall that applies to objects moving vertically under the influence of gravity, the y-axis has been traditionally adopted. It is just more indicative of the motion than the x-axis.

2. In the spirit of Remark 1, in the following are two sets of equations, each of which embraces only one axis.

One-dimensional motion along the x-axis

$$v = v_o + at, \tag{2.14a}$$

$$x = v_o t + \frac{1}{2}at^2, \tag{2.14b}$$

$$v^2 = v_o^2 + 2\,ax, \tag{2.14c}$$

$$x = \bar{v}t. \tag{2.14d}$$

One-dimensional motion along the y-axis

$$v = v_o + at, \tag{2.15a}$$

$$y = v_o t + \frac{1}{2}at^2, \tag{2.15b}$$

$$v^2 = v_o^2 + 2\,ay, \tag{2.15c}$$

$$y = \bar{v}t. \tag{2.15d}$$

In the above equations, v_o is understood to be along the appropriate axis.

2.7 FREE-FALLING OBJECT

This section is devoted to motion of objects in a vertical direction that are subject to the Earth's gravitational acceleration. The Earth's gravitational acceleration, near its surface, is considered constant. Its value is 9.80 m/s^2 and it is always directed toward the center of the Earth. The motion of an object under gravitational acceleration is a good example of one-dimensional motion. Since the motion is vertical, it would be appropriate to use the y-axis in describing it. The y-axis will be taken as positive upward (Figure 2.5) and all vectors such as displacement and velocity that are directed along the positive y-axis will be considered positive, whereas those directed along the negative y-axis will be considered negative.

One important caution, here, pertains to the acceleration of any object moving freely under gravitation. Since this acceleration is always directed downward, it will be considered negative, regardless of whether the direction of the motion of the object is upward or downward. Thus, in Equation 2.15

$$a = -9.80\ m/s^2. \tag{2.16}$$

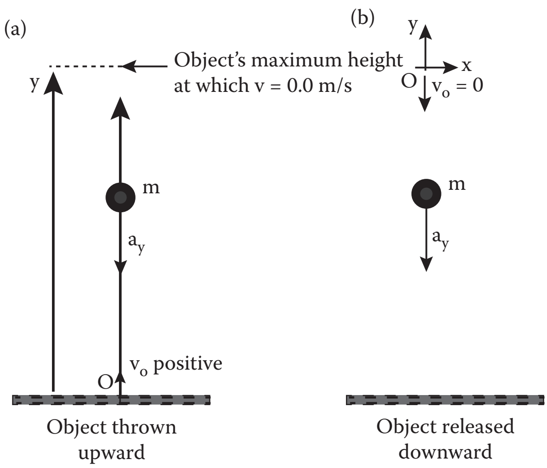


FIGURE 2.5 Motions of a free-falling object: (a) the object is thrown upward and (b) the object is released downward. In both cases, the y-axis is positively directed upward, whereas the acceleration of the object is 9.80 m/s^2 downward.

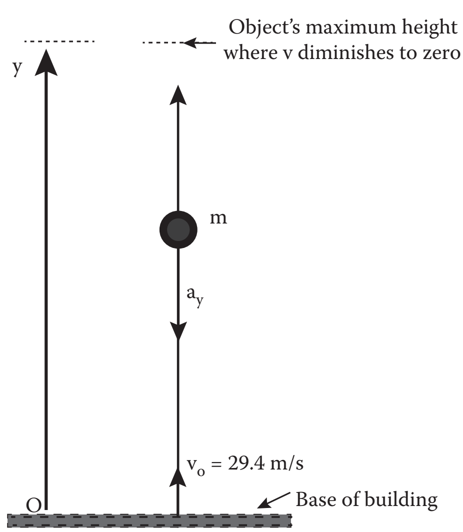
2.7.1 Important Practical Remarks

1. In any motion, take the origin at the location where the motion starts.
2. Accordingly, at that initial position, x = 0 if the motion is in the x direction and y = 0 if the motion is in the y direction.
3. In both cases above, t = 0 at the origin, O.

EXAMPLE 2.8

A ball is projected upward from the base of a building with a velocity $v_o = 29.4$ m/s (the figure below). If air resistance is ignored, find

a. How much time it takes for the ball to get to its maximum height
b. Its maximum height



Solution

a. Since v_o and a are known, to find the time of the ball's flight upward, use the equation of motion

$$v = v_o + at,$$

and substituting $v_o = 29.4$ m/s, $a = -9.80$ m/s^2, and $v = 0.00$ m/s at the highest point gives

$$0\ \text{m/s} = 29.4\ \text{m/s} - (9.80\ \text{m/s}^2)\ t.$$

Solving for t gives

$$t = 3.00\ \text{s}.$$

b. For the maximum height that the ball reaches, use the equation

$$y = v_o t + \frac{1}{2}at^2.$$

Substituting for all the known quantities gives

$$y = (29.4\ \text{m/s})(3.00\ \text{s}) - \frac{1}{2}(9.80\ \text{m/s}^2)(3.00\ \text{s})^2 = 88.2 - 44.1 = 44.1\,\text{m}.$$

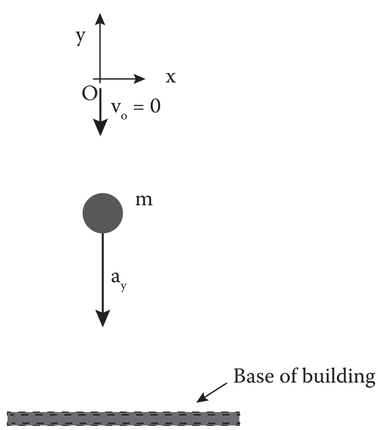
Analysis

Note that the acceleration *a* as substituted above was -9.80 m/s^2 and not 9.80 m/s^2.

EXAMPLE 2.9

A ball is released off a balcony on the 10th floor of a building, 44.1 m above the ground (the figure below). If air resistance is considered negligible, find

a. How much time does it take the ball to get to the ground
b. The value of the vertical velocity of the ball as it hits the ground



Solution

a. Since y, v_o, and a are known, this suggests using Equation 2.15b. Upon substituting for

$$y = -44.1 \text{ m}, \quad v_o = 0.00 \text{ m/s}, \quad \text{and} \quad a = -9.80 \text{ m/s}^2,$$

the above equation becomes

$$-44.1 = (0)\,t - \frac{1}{2}(9.80)\,(t)^2.$$

which after basic algebra gives

$$t = 3.00 \text{ s}.$$

b. Since v_o, a, and t are now known, using the equation $v = v_o + at$ and substituting for

$$v_o = 0.00 \text{ m/s}, \quad t = 3.00 \text{ s}, \quad \text{and} \quad a = -9.80 \text{ m/s}^2$$

gives

$$v = 0 - (9.80)(3.00) = -29.4 \text{ m/s}.$$

Thus, the velocity of the ball as it hits the ground is –29.4 m/s.

The negative velocity indicates a downward direction, that is, opposite to the positive direction of the y-axis adopted upward. This is consistent with the earlier remark.

Analysis

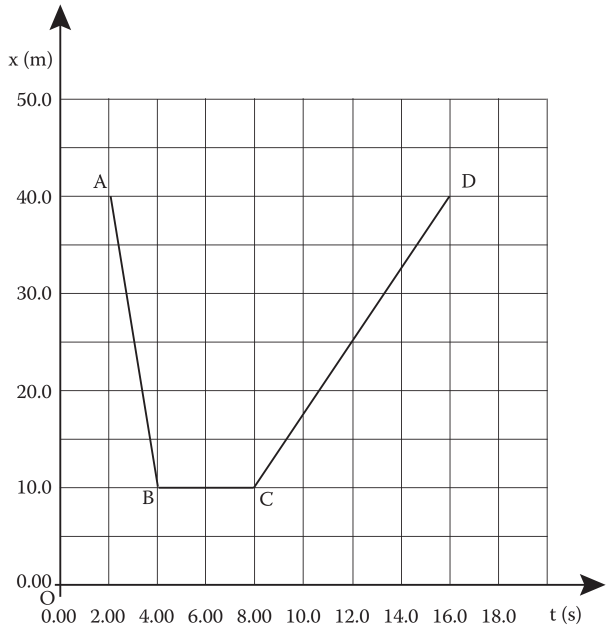
1. From part (a), it was found that an object projected upward with an initial velocity v_o gets to its highest point in time t equal to the time elapsed for the object to fall back to the ground. In other words, the round trip time of the projectile's flight is twice the time of the flight upward.
2. In making a connection between Examples 2.8 and 2.9, we can say that the object returns to the ground with a final velocity, just prior to hitting the ground, equal to its initial velocity with which it was launched vertically upward.
3. Please note that in Equation 2.15b, y is the displacement of the ball after some time t = 3.00 s. Since the ground is below the release point on the balcony, the ground coordinate is negative with respect to the balcony where the zero of the motion is taken. This means that the balcony is 44.1 m above the ground.

PROBLEMS

2.1 A bicycler starting from Indiana, Pennsylvania, makes three successive displacements; she moves to location A, 725 m east, then moves to location B, 1250 m west of A, and finally moves east again to location C with a distance of 775 m from B. If east is chosen as the positive direction, what
   a. Is the bicycler's resultant displacement?
   b. Is the distance traveled by the bicycler?

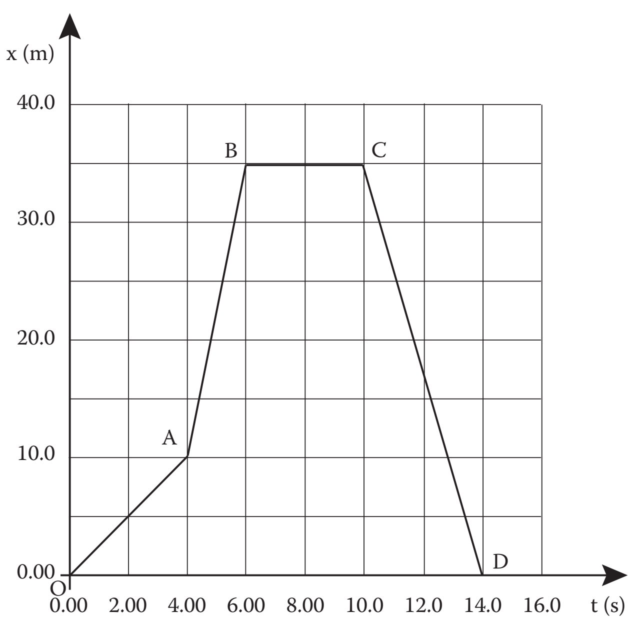
2.2 In the previous problem, if the total time it took the bicycler in this trip is 22 min, determine
   a. The bicycler's average velocity
   b. The bicycler's speed

2.3 The following figure displays the displacement x versus time t for an object moving in x direction. Use the depicted graph to calculate
   a. The displacement of the object in the interval between t = 2.00 s and t = 8.00 s.
   b. The displacement of the object in the interval between t = 2.00 s and t = 16.0 s.
   c. The average velocity of the object in the regions AC and AD.

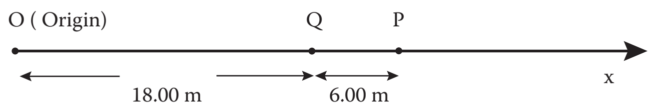


2.4 Using the data and graph in the previous problem, calculate
   a. The instantaneous velocity of the object at t = 3.0 s, t = 6.0 s, and t = 11 s.
   b. The average speed of the object in its whole motion between A and D.

2.5 The figure below displays the position x versus time t for an object moving in the x direction. Use the x–t graph to calculate
   a. The displacement of the object in the following intervals: 0.00–4.00 s, 4.00–6.00 s, and 6.00–14.00 s.
   b. The average velocity of the object in the regions OA, AB, and BD.



2.6 A soccer player, while training for the Olympics, moves along a straight line, taken as the x-axis, from point O to P in 6.0 s and back to Q in 2.0 s (the figure below). Using the distances illustrated in the diagram, determine
 a. The average velocity of the player between O and P and between P and Q.
 b. The average velocity of the player between her initial and final positions.
 c. The average speed of the player during her motion between her initial and final positions.



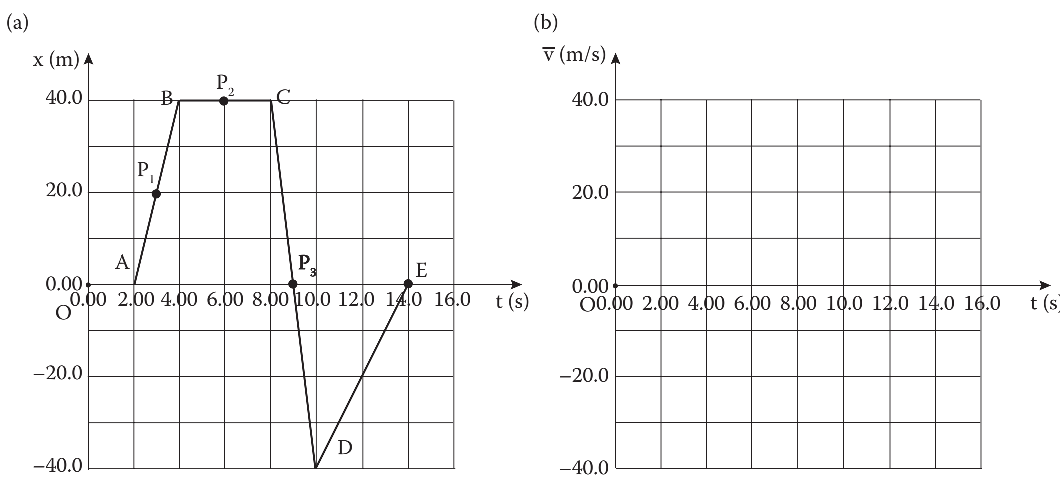
2.7 A motorist planned to join a group gathering at a site 222 km away east. He covered the first 111 km in 80.0 min and the remaining distance in 40.0 min. Find
 a. The average velocity of the motorist during this trip
 b. The average speed of the motorist during this trip

2.8 A driver is driving her car between two cities, Indiana and the suburb of Philadelphia, 416 km apart. She drove the first 226 km with an average velocity $\bar{v}_1 = 62.0$ km/h, and after stopping for gas for 15 min drove the remaining distance with an average velocity $\bar{v}_2 = 68.0$ km/h. Assuming the highway was along a straight west–east turnpike, determine
 a. The total time that took the driver to get to her destination
 b. The average velocity of the driver during her whole trip
 c. The average speed of the driver during her whole trip

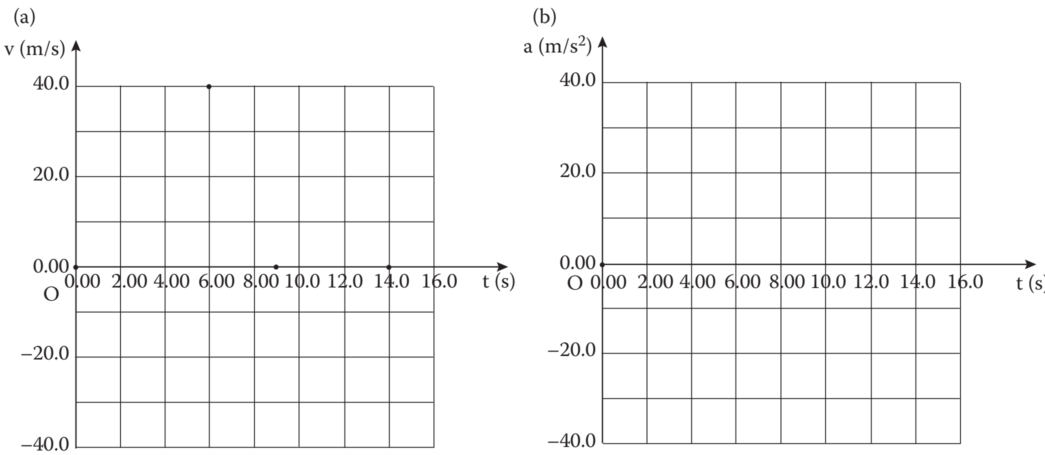
2.9 A basketball player can jump 1.20 m off a hardwood floor. With what upward velocity did he leave the floor?

2.10 A ball is thrown downward from the balcony of a third-floor apartment with an initial velocity of 8.00 m/s. A stop watch measures the ball's time from the balcony to the ground to be 3.20 s. If air resistance is assumed negligible
 a. How high is the balcony above the ground?
 b. Determine the velocity of the ball just before it reaches the ground.

2.11 The figure below (part (a)) shows the displacement x versus time t for an object moving in the x direction. Use the x–t graph to calculate the average velocity of the object in the following regions: AB, BC, D, and DE. Sketch the plots of v versus t in these regions and display them in part (b) that is also provided for you.



2.12 Use the figure above (part (a))
   a. To determine the instantaneous velocity versus t of the object in the middle of each of the regions AB, BC, and CD
   b. To determine the average acceleration of the object in the regions AB, BC, and CD
   c. Then sketch the v–t and the a–t plots for the above regions and display them on the charts provided in the figure below (part (a) and (b)), respectively.



2.13 In a nappy instant of a tired student in a physics class, his pencil fell from his hand, 0.490 m above the floor. Rounding off your answer to three significant figures, determine
   a. The time it took for the pencil to reach the floor
   b. The velocity of the pencil just before it hits the floor
   c. The average velocity of the pencil during its falling period of time

2.14 A stone of a mass m = 0.220 kg is thrown vertically upward from the ground with an initial velocity v_o = 14.0 m/s. Ignoring air resistance, determine
   a. The maximum height above the ground the stone reaches
   b. The time it takes the stone in its round trip motion

2.15 A stone is thrown upward from a balcony, 15.0 m above the ground, with a velocity v_o = 21.0 m/s. Ignoring air resistance, determine
   a. The maximum height above the balcony the stone reaches
   b. The time duration of the stone in air before it hits the ground

3 Two-Dimensional Motion and Circular Motion

As seen for motions in one dimension (Chapter 2), the quantities displacement, velocity, and acceleration were allowed to change along one direction, the x-direction or the y-direction. To fulfill the vector aspect of these quantities as they change along one of the coordinate axes, say x, it was sufficient to have them assume positive or negative values. In other words, the sign of these quantities was indicative enough of the direction along which the quantity is pointing. Motion in two dimensions, however, is more involved and will require the use of vectors.

In two-dimensional motion, the quantities displacement, velocity, and acceleration can be resolved into two components. If the plane of motion is taken as the x–y plane, one of the two components can be chosen to be along the x-axis and the other component can be chosen to be along the y-axis. As is known, a Cartesian coordinate system consists of mutually perpendicular independent axes, which imply that all vector components along these axes can be treated separately. As will be seen, this important feature of Cartesian coordinates tremendously simplifies our handling of the kinematics of motion in two dimensions. This feature in fact reduces the problem to constituent problems, each of which is in one direction. A full description of the two-dimensional motion can then be formed from its component motions. The only parameter that is common to these two motions is time that, at any instant of an object's motion, occurs in both x and y motions simultaneously.

3.1 DISPLACEMENT, VELOCITY, AND ACCELERATION

The displacement, $\Delta\mathbf{r}$, of an object is a vector quantity that describes a change in position of the object from a position vector, $\mathbf{r}_1$, to another position, $\mathbf{r}_2$ (Figure 3.1). It can be seen that $\Delta\mathbf{r}$ consists of two vector components, $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$. Hence, it can be expressed as

$$\Delta\mathbf{r} = \Delta\mathbf{x} + \Delta\mathbf{y}. \tag{3.1}$$

From the above definition, the average velocity $\bar{\mathbf{v}}$ along the displacement $\Delta\mathbf{r}$ can be obtained simply by dividing the above equation by Δt. This process gives

$$\bar{\mathbf{v}} = \bar{\mathbf{v}}_x + \bar{\mathbf{v}}_y, \tag{3.2}$$

where $\bar{\mathbf{v}}_x$ and $\bar{\mathbf{v}}_y$ are the average velocity components along the x and y axes, respectively. The instantaneous velocity $\mathbf{v}$, however, of the object takes a generalized form of that introduced in the one-dimensional motion (see Chapter 2). Thus, the instantaneous velocity $\mathbf{v}$ in two dimensions is

$$\mathbf{v} = \lim_{\Delta t \to 0}\left(\frac{\Delta\mathbf{r}}{\Delta t}\right). \tag{3.3}$$

Of course, each of the two components, v_x and v_y, follows the basic definition of a one-dimensional motion. In other words

$$v_x = \lim_{\Delta t \to 0}\left(\frac{\Delta x}{\Delta t}\right), \tag{3.4a}$$

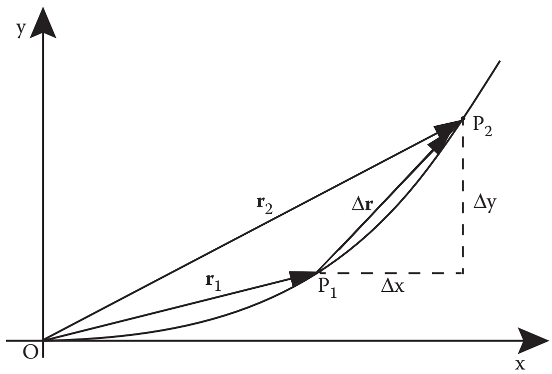


FIGURE 3.1 An object moving between points P_1 and P_2 that are described by the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively. $\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ is the displacement vector between the two points.

$$v_y = \lim_{\Delta t \to 0}\left(\frac{\Delta y}{\Delta t}\right). \tag{3.4b}$$

By the same token, the acceleration, **a**, in two dimensions has two components a_x and a_y. Again, the basic definition of **a** in two dimensions is analogous to that introduced for the acceleration in one-dimensional motion. It takes the general form

$$\mathbf{a} = \frac{\mathbf{v}_2 - \mathbf{v}_1}{t_2 - t_1}.$$

The two components a_x and a_y of **a** should follow the same patterned definition, that is

$$a_x = \frac{v_x - v_{ox}}{t}, \tag{3.5a}$$

$$a_y = \frac{v_y - v_{oy}}{t}. \tag{3.5b}$$

In the above equations, v_x and v_y are the final velocity along x and y, respectively, and v_{ox} and v_{oy} are the initial velocities along x and y, respectively; the interval $t_2 - t_1$ is written as $t - 0 = t$.

Again, similar to our argument in Chapter 2, the acceleration in either direction could be either positive, implying that the object in that direction is accelerating, or negative, implying that the object in that direction is decelerating.

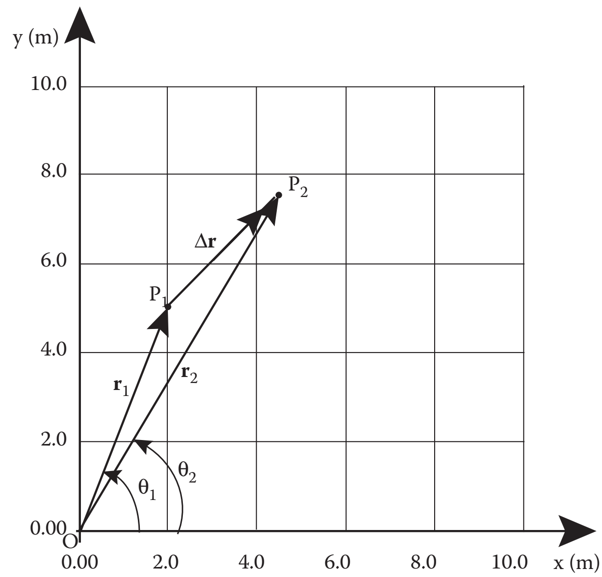
It is also quite important to stress that in all cases of the two-dimensional motion treated here, the acceleration is constant in magnitude and direction. Accordingly, each of the two components a_x and a_y will always be independently constant.

EXAMPLE 3.1

An object has moved from position P_1 (2.00, 5.00 m) to position P_2 (4.50, 7.50 m) (the following figure). Determine the magnitude and direction of

a. The initial position vector

b. The final position vector
c. The displacement vector



Solution

a. The magnitude of the initial position $\mathbf{r}_1$ vector (see Equation 1.9) is

$$r_1 = \sqrt{x_1^2 + y_1^2} = \sqrt{(2.00\text{ m})^2 + (5.00\text{ m})^2} = 5.40\text{ m}.$$

The direction θ_1 of the initial position $\mathbf{r}_1$ vector (see Equation 1.10) is

$$\theta_1 = \tan^{-1}\left(\frac{y_1}{x_1}\right) = \tan^{-1} = \left(\frac{5.00\text{ m}}{2.00\text{ m}}\right) = \tan^{-1}(2.50) = 68.2°.$$

b. The magnitude and direction of the final position $\mathbf{r_2}$ vector are similary calculated.

$$r_2 = \sqrt{x_2^2 + y_2^2} = \sqrt{(4.50\text{ m})^2 + (7.50\text{ m})^2} = 8.70\text{ m}.$$

$$\theta_2 = \tan^{-1}\left(\frac{y_2}{x_2}\right) = \tan^{-1} = \left(\frac{7.50\text{ m}}{4.50\text{ m}}\right) = \tan^{-1}(1.67) = 59.1°.$$

c. The displacement vector $\Delta\mathbf{r}$ has two components Δx and Δy. Numerically

$$\Delta x = x_2 - x_1 = 4.50 \ - 2.00\text{ m} = 2.50\text{ m}.$$

$$\Delta y = y_2 - y_1 = 7.50 \ - 5.00\text{ m} = 2.50\text{ m}.$$

The magnitude Δr and direction θ of the displacement vector **Δr** are given by
Magnitude

$$\Delta r = \sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{(2.50\ \text{m})^2 + (2.50\ \text{m})^2} = 3.56\ \text{m}.$$

Direction

$$\theta = \tan^{-1}\left(\frac{\Delta y}{\Delta x}\right) = \tan^{-1} = \left(\frac{2.50\ \text{m}}{2.50\ \text{m}}\right) = \tan^{-1}(1.00) = 45.0°.$$

Analysis

1. From the above values in parts (a), (b), and (c), one notices that the directions of the initial position vector, final position vector, and the displacement vector are all different; in each case, the obtained angle is the angle that the vector makes with the positive x-axis.
2. As an exercise, one could draw the three vectors: the initial position vectors, final position vector, and the displacement vector, and check whether the above analytical solution is consistent with a graphical solution of the vector equation $\Delta \mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ (see Example 1.14).

EXAMPLE 3.2

Knowing that the displacement in Example 3.1 has occurred in 2.00 s, what would be

a. The x-component of the object's average velocity?
b. The y-component of the object's average velocity?
c. The magnitude and direction of the object's average velocity?

Solution

a. As Δx from the previous example is Δx = 2.50 m and t, that is, Δt = 2.00 s, then

$$\bar{\mathbf{v}}_x = \frac{\Delta x}{\Delta t} = \frac{2.50\ \text{m}}{2.00\ \text{s}} = 1.25\ \text{m/s}.$$

b. $\bar{\mathbf{v}}_y = \dfrac{\Delta y}{\Delta t} = \dfrac{2.50\ \text{m}}{2.00\ \text{s}} = 1.25\ \text{m/s}.$

c. Thus, the magnitude of the object's average velocity is

$$\bar{v} = \sqrt{(\bar{v}_x)^2 + (\bar{v}_y)^2} = \sqrt{(1.25\ \text{m/s})^2 + (1.25\ \text{m/s})^2} = 1.77\ \text{m/s}.$$

The direction of the object's average velocity is

$$\theta = \tan^{-1}\left(\frac{\bar{v}_y}{\bar{v}_x}\right) = \tan^{-1}\left(\frac{1.25\ \text{m/s}}{1.25\ \text{m/s}}\right) = 45.0°.$$

Analysis

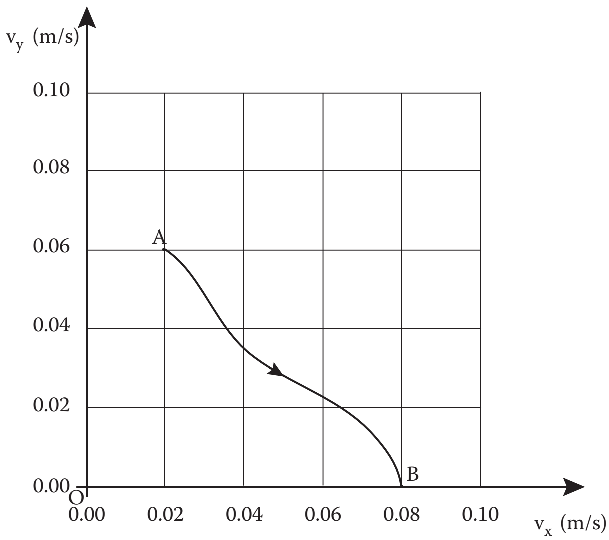
Notice that the direction of the average velocity is the same as the direction of the displacement found in the previous example. Since the average velocity vector is defined as $\bar{\mathbf{v}} = (\Delta\mathbf{r}/\Delta t)$, it is quite expected that the displacement vector dictates the direction of the average velocity.

The following example gives information about the instantaneous velocities of an object at its initial and final positions; so, one can determine the object's acceleration as it moves between the two positions.

EXAMPLE 3.3

A spider is crawling from point A to point B along the path shown in the figure below in 4.0 s. If the spider's velocity components at point A are $v_x = 0.02$ m/s and $v_y = 0.06$ m/s and the velocity components at B are $v_x = 0.08$ m/s and $v_y = 0.00$ m/s, find

a. The components of the spider's average acceleration a_x and a_y between A and B.
b. The magnitude of the spider's acceleration between these two points.



Solution

a. $\bar{a}_x = \Delta v_x/\Delta t = (0.08 \text{ m/s} - 0.02 \text{ m/s})/(4.0 \text{ s}) = 0.02 \text{ m/s}^2$.

$\bar{a}_y = \Delta v_y/\Delta t = (0.00 \text{ m/s} - 0.06 \text{ m/s})/(4.0 \text{ s}) = -0.02 \text{ m/s}^2$.

b. $\bar{a} = \sqrt{\bar{a}_x^2 + \bar{a}_y^2} = \sqrt{(0.02 \text{ m/s}^2)^2 + (-0.02 \text{ m/s}^2)^2} = 0.03 \text{ m/s}^2$.

Analysis

Note that all answers were rounded off to one significant figure because the values in the plot can be read to only one significant figure.

3.2 EQUATIONS OF MOTION IN TWO DIMENSIONS

The argument that leads to the derivation of the equations of motion in one dimension applies to each of the components of motion in two-dimensional motion, and since each component of this motion is independent of the other component, two sets of equations, one for each component, can be independently applied. Accordingly, the following two sets of equations should provide a full description of an object's two-dimensional motion.

The x-motion

$$v_x = v_{ox} + a_x t, \tag{3.6a}$$

$$x = v_{ox}t + \frac{1}{2}a_x t^2, \tag{3.6b}$$

$$v_x^2 = v_{ox}^2 + 2a_x x, \tag{3.6c}$$

$$\bar{v}_x = \frac{1}{2}(v_x + v_{ox}). \tag{3.6d}$$

The y-motion

$$v_y = v_{oy} + a_y t, \tag{3.7a}$$

$$y = v_{oy}t + \frac{1}{2}a_y t^2, \tag{3.7b}$$

$$v_y^2 = v_{oy}^2 + 2\,a_y y, \tag{3.7c}$$

$$\bar{v}_y = \frac{1}{2}(v_y + v_{oy}). \tag{3.7d}$$

3.2.1 Important Remarks

1. It is often convenient to take the origin of the coordinate system at the location where the motion starts.
2. If Remark 1 is followed, then for motion in the x–y plane, $x = 0$ and $y = 0$ would mark the initial positions of the object when $t = 0$.
3. Solving Equations 3.6 and 3.7 for x and y at any given instant t would define the object positions at time t.

3.3 PROJECTILE MOTION

In Chapter 2, the free fall of an object under gravity was treated. It was a one-dimensional vertical motion, and hence the y coordinate was a fit to describe it. In a projectile motion where an object is projected with an angle θ_o with the horizontal, the object goes up and down (vertical motion) and moves away (horizontal motion) from the launching position as well. This kind of motion is demonstrated in many circumstances such as when playing football, baseball, firing a canon in a battlefield, or launching an Earth-to-Earth rocket; in fact, such projectile motions are real examples of two-dimensional motion. The vertical motion of the object is subjected to the downward gravitational acceleraion g. However, along the horizontal motion, there is no acceleration backward nor forward (air resistance is ignored). Adopting two axes, x and y such that the positive y-axis is directed upward and the positive x-axis is directed forward toward the horizontal movement of the object, we can then set $a_x = 0$ and $a_y = -g$. Figure 3.2 is a typical illustration of projectile motion.

In the projectile motion in Figure 3.2, the object could be a ball that is kicked with an angle θ_o with the horizontal. The ball moves away horizontally until hitting the ground at a landing position, L. The other component of the ball's motion is the vertical motion, moving upward until it reaches a maximum height h, at point H, after which the ball starts to fall downward. So, the rising of the ball occurs in segment OH of the path and the ball's falling occurs in segment HL of the path. The ball's horizontal displacement between O and L can be described as a displacement along the x-axis,

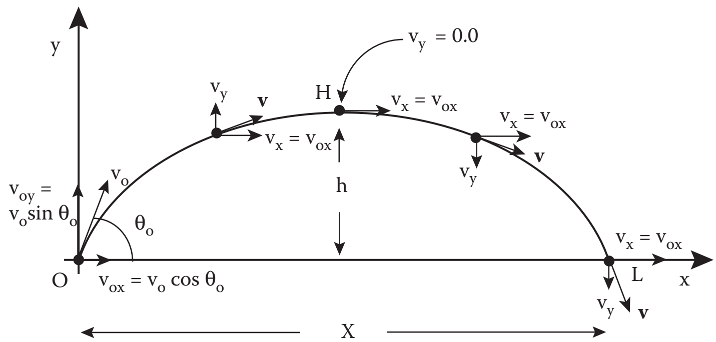


FIGURE 3.2 A trajectory of a projectile with several of its locations and velocity vectors along the trajectory.

and the ball's vertical displacement from O to H and back to L can be described by corresponding displacements along the y-axis.

Analyzing this motion can start with resolving the initial velocity, $\mathbf{v}_o$, into two components, an x-component $v_{ox} = v_o \cos \theta_o$ along the x-axis, and a y-component $v_{oy} = v_o \sin \theta_o$ along the y-axis. The reality that the x component of the acceleration of the projectile is zero means that the projectile's velocity along the x-axis experiences no change. Accordingly, the projectile's velocity along the x-axis at any instant during the projectile's motion remains equal to its initial value, that is, $v_x = v_{ox} = v_o \cos \theta_o$.

In light of the above discussion, the two sets of equations describing the two-component motions of the projectile can be summarized in Table 3.1.

EXAMPLE 3.4

A ball is rolled off a table's top (the following figure) with an initial horizontal speed of 5.96 m/s. A stopwatch measures the ball's trajectory time from the table's top to the floor to be 0.300 s. (air resistance is assumed negligible).

a. Determine the height of the table.
b. How far away from the table does the ball land?
c. Determine the vertical velocity of the ball as it hits the ground.
d. Determine the horizontal velocity of the ball as it hits the ground.

TABLE 3.1
Summary of the Two Cartesian Components of a Projectile

Projectile Motion (Motion in the x–y Plane)			
x-Motion: $a_x = 0.00$		**y-Motion: $a_y = -g = -9.80\ m/s^2$**	
$v_x = v_{ox} = v_o \cos \theta_o$	(3.8a)	$v_y = v_{oy} - gt = v_o \sin \theta_o - gt$	(3.9a)
$x = v_{ox}\, t = (v_o \cos \theta_o)t$	(3.8b)	$y = v_{oy}t - \frac{1}{2}gt^2 = (v_o \sin \theta_o)t - \frac{1}{2}gt^2$	(3.9b)
$v_x^2 = v_{ox}^2,\ v_x = v_{ox}$	(3.8c)	$v_y^3 = v_{oy}^2 - 2gy = (v_o \sin \theta_o)^2 - 2gy$	(3.9c)

Note: x is the horizontal direction along the ground and y is the vertical direction above the ground.

Solution

a. If $y = v_{oy}\,t - (1/2)gt^2$, then upon substituting for the known quantities, this becomes

$$y = 0.00 \text{ m/s(t)} - \frac{1}{2}(9.80 \text{ m/s}^2)(0.300 \text{ s})^2.$$

So

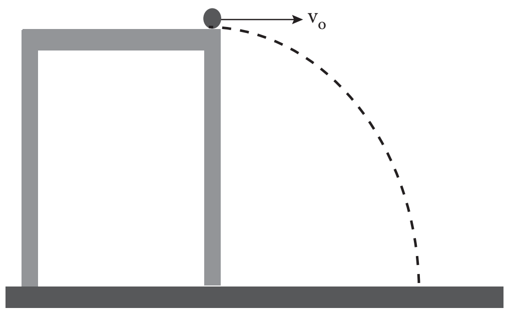
$$y = -0.440 \text{ m}.$$

That is, the height of the table is 0.440 m above the floor.

b. The range $X = v_{ox}t = (5.96 \text{ m/s})(0.300 \text{ s}) = 1.79$ m.

c. $v_y = v_{oy} - gt = 0 - (9.80 \text{ m/s}^2)\ (0.300 \text{ s}) = -2.94$ m/s.

d. Since $v_x = v_{ox}$, then $v_x = 5.96$ m/s; this does not change in a projectile motion.



Analysis

1. Notice that since the ball was rolled off the table, it had an initial velocity only in the x-direction. Its initial velocity in the vertical direction, that is, in the y direction, was zero. Since the horizontal and y components of this motion are independent, the vertical component of this motion will not be different from an object released from the same height as the table.
2. The only common quantity between the x and y components of this motion is the time, t. The object needs as much time to move down 0.440 m as it needs to hit the ground along the horizontal direction, 1.79 m away from the table.

Exercise

One can solve the one-dimensional vertical motion of the above exercise and find the height of the table from observing the time of fall down of the ball from the edge of the table to be 0.300 s.

EXAMPLE 3.5

A soccer ball is kicked with an initial velocity of 20.0 m/s making an angle of 37° with the horizontal (the following figure). Determine

a. How long it takes the ball to hit the ground
b. How high the ball rises
c. How far from the player the ball hits the ground

Solution

a. First resolve the initial velocity of the ball into two components, the horizontal x-component and the vertical y-component.

$$v_{ox} = (v_o \cos \theta_o) = (20.0 \text{ m/s})(0.80) = 16 \text{ m/s},$$

$$v_{oy} = (v_o \sin \theta_o) = (20.0 \text{ m/s})(0.60) = 12 \text{ m/s}.$$

To find the time of flight, first find the time it takes the ball to reach its maximum height at which $v_y = 0$. For that, use the equation of motion

$$v_y = v_{oy} - gt_{top}.$$

Substituting for $v_y = 0$ and $g = 9.80 \text{ m/s}^2$ gives

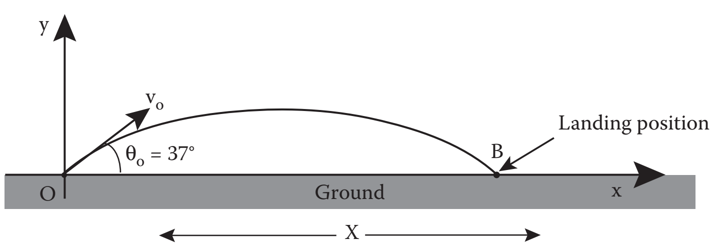
$$0.00 \text{ m/s} = 12 \text{ m/s} - (9.80 \text{ m/s}^2)t_{top},$$

from which

$$t_{top} = 1.2 \text{ s}.$$

Thus, since the time of the whole flight is twice the time the ball takes during its upward motion (see Example 2.9), the time of flight is

$$t = 2(t_{top}) = 2(1.2) = 2.4 \text{ s}.$$



b. Since the time it takes the ball to get to its maximum height is now known, using the second equation of motion

$$y = v_{oy}t - \frac{1}{2}gt^2,$$

and substituting for all given quantities yields y at the highest point of the ball's path as

$$y_{\text{highest point}} = (12 \text{ m/s})(1.2 \text{ s}) - \frac{1}{2}(9.80 \text{ m/s}^2)(1.2 \text{ s})^2.$$

Thus,

$$y_{\text{highest point}} = 14 \text{ m} - 7.1 \text{ m} = 7 \text{ m}.$$

c. Since the horizontal x component of the ball's velocity v_x is constant, the horizontal displacement of the ball X is

$$X = v_{ox}\, t = (16\ \text{m/s})(2.4\ \text{s}) = 38\ \text{m}.$$

Analysis

Once the time to reach its maximum height is known, insertion in the x and y components of the ball's initial velocity makes the solution rather simple. Each component of motion is then treated as one-dimensional motion.

3.4 UNIFORM CIRCULAR MOTION

Two kinds of motion have been treated so far, one- and two-dimensional motion. We started with the simplest kind of motion, the motion along one line. The second was two-dimensional motion, which is more involved, but still carries a lot of simplification of an object's real motion. The more realistic two-dimensional motion of an object is generally along a curve that could be circular, elliptical, or of some other arbitrary shape. The extremely interesting features and properties of circular motion or near-circular motion (electrons in their atomic orbits, planets in their orbits, cars driven on flat curved roads, ramps, etc.) warrants special attention and extended treatment. The simplest circular motion is known as uniform circular motion, where an object rotates in a fixed circle with a constant speed. The conditions that make this rotation possible will be discussed in the next chapter.

This motion is introduced with the general definition of the average acceleration that an object is subjected to during a time interval Δt. Assumed to be constant, this was stated for one- and two-dimensional motion in the general form

$$\mathbf{a} = \frac{\Delta \mathbf{v}}{\Delta t}. \tag{3.10}$$

Since **v** is a vector quantity that has a magnitude and direction, the above definition may involve a change in the magnitude of the object's velocity, **v**, a change in the direction of its velocity, or both. Accordingly, an object executing motion that is not along a straight line is expected to have an acceleration even if it is moving with a velocity of constant magnitude because this velocity is constantly changing in direction. This is the essence of uniform circular motion, in which an object is rotating along a circular path with a velocity **v**, also known as the tangential velocity, which is of a fixed magnitude but constantly changing in direction. The notion for **v** is tangential because the direction of this velocity is along the tangent to the circular path. Thus, the object experiences an acceleration resulting from the continuous change in the direction of **v**, along the circular path. This acceleration is always directed toward the center and is thus called centripetal acceleration.

To illustrate this problem, consider an object rotating on a circle with a constant speed, v. Take two positions, P_1 and P_2 along the object's path on the circle (Figure 3.3a). The position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ describe the two positions of the object at P_1 and P_2, respectively, relative to the center of the circle, C. Assume that the angle between the two position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ is φ. Of course, the magnitudes of the radii $\mathbf{r}_1$ and $\mathbf{r}_2$ are equal. The velocities $\mathbf{v}_1$ and $\mathbf{v}_2$ of the object at these points have different directions, but are of the same magnitude. Therefore, we may resketch these two velocity vectors as in Figure 3.3b, preserving their magnitudes and directions (see Section 1.7). Visualizing the object when the interval Δt becomes smaller, point P_2 becomes closer to P_1, making the change in the displacement vector $\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ smaller. And since the angle between v_1 and v_2 is equal to the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$, the change in the velocity vector would also become smaller (Figure 3.3c). As the time interval becomes infinitesimally small, the direction of $\mathbf{v}_2$ starts to get more aligned along $\mathbf{v}_1$ and the change in the velocity vector $\Delta\mathbf{v} = \mathbf{v}_2 - \mathbf{v}_1$ becomes infinitesimally small (Figure 3.3d) such that it becomes perpendicular to $\mathbf{v}_1$. In Figure 3.3a, since $\mathbf{v}_1$ is perpendicular to the radius

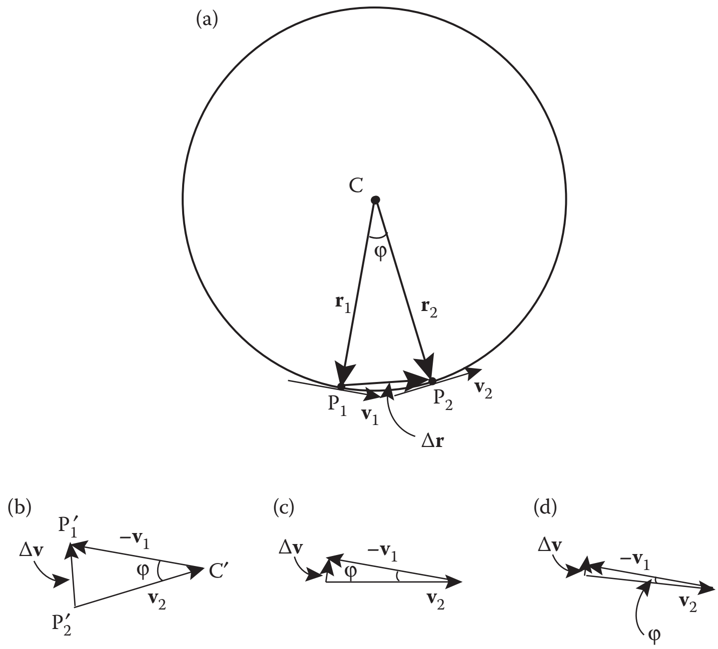


FIGURE 3.3 An object rotating on a circle with a velocity of constant value, v. (a) Two positions P_1 and P_2 at which the object's velocities are $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively; (b), (c), and (d) cases of progressive changes in the time interval Δt. In case (d), $\Delta t \to 0$, which corresponds to making the change in the velocity vector $\Delta\mathbf{v}$ almost aligned along r_1, pointing toward C.

r_1 when Δt approaches zero, both $\Delta\mathbf{v}$ and r_1 become perpendicular to the same vector $\mathbf{v}_1$. That is, $\Delta\mathbf{v}$ and $\mathbf{r}_1$ are parallel, with $\Delta\mathbf{v}$ pointing opposite to $\mathbf{r}_1$. In other words, $\Delta\mathbf{v}$ and $\mathbf{r}_1$ are antiparallel. This means that the vector $\Delta\mathbf{v}$ points toward the center of the circle, C, instead of pointing outward.

The limiting case is demonstrated separately in Figure 3.4. The arrow representing the change in the object's velocity $\Delta\mathbf{v}$ was laid side by side to the radius vector, $\mathbf{r}_1$, and not on it just for clarity and to avoid confusion. Since the acceleration just defined should be in the direction of $\Delta\mathbf{v}$, it can be concluded that this acceleration actually points toward the center of the circle along which the object is revolving. Notice that each of the triangles in (b), (c), and (d) has two sides that are equal in magnitude $v_1 = v_2 = v$. The angle between them is also φ equal to that between $\mathbf{r}_1$ and $\mathbf{r}_2$, case (a). In a parallel notation used for the triangle CP_1P_2 in (a), the triangle in case (b) is labeled as $C'P_1'P_2'$.

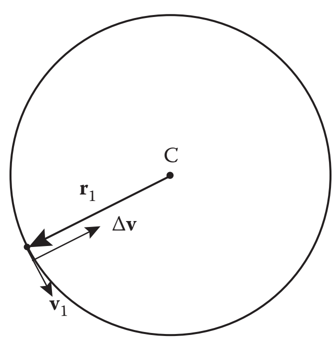


FIGURE 3.4 In a uniform circular motion, an infinitesimally small interval Δt between two positions P_1 and P_2 brings the points too close to each other such that the direction of $\Delta\mathbf{v}$ at P_1 becomes perpendicular to $\mathbf{v}_1$, that is, antiparallel to the radius position vector $\mathbf{r}_1$.

As for the magnitude of this centripetal acceleration, it can be derived upon considering the two triangles CP_1P_2, consisting of sides r_1, r_2, and Δr, and $C'P_1'P_2'$, consisting of sides v_1, v_2, and Δv, in Figure 3.5. The triangle $C'P_1'P_2'$ is rotated 90° clockwise just to make the geometry more indicative. As the two triangles are similar, the following ratios $\frac{\text{base in triangle (a)}}{\text{base in triangle (b)}}$ and $\frac{\text{side in triangle (a)}}{\text{side in triangle (b)}}$ are equal. That is

$$\frac{\Delta r}{\Delta v} = \frac{r}{v},$$

or

$$\Delta v = \left(\frac{v}{r}\right)\Delta r.$$

Thus,

$$\left(\frac{\Delta v}{\Delta t}\right) = \left(\frac{v}{r}\right)\left(\frac{\Delta r}{t}\right)$$

or

$$\left(\frac{\Delta v}{\Delta t}\right) = \left(\frac{v}{r}\right)v.$$

That is

$$a_c = \frac{v^2}{r}. \tag{3.11}$$

And as just stated, this acceleration is in a direction pointing toward the center of the circle. That is why this acceleration acquires the notion "centripetal." It is worth noting that the centripetal acceleration is also called radial acceleration. Figure 3.6 shows the acceleration of such an object at three points P, Q, and R, rotating with a velocity v along the circumference of a circle of radius r and center C.

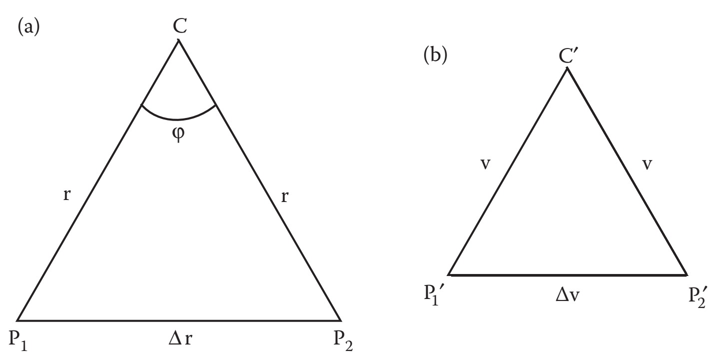


FIGURE 3.5 Triangles in (a) and (b) are similar.

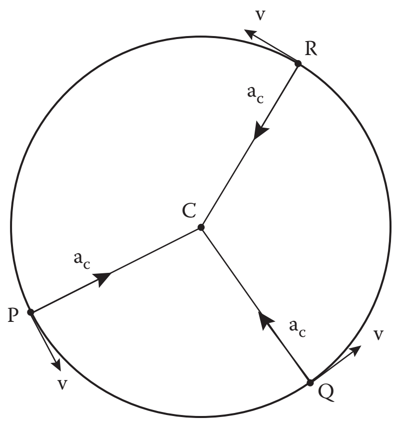


FIGURE 3.6 An object whose velocity along the circumference of a circle has the same value experiences a centripetal acceleration a_c at all points of its circulation.

In Equation 3.11, notice that a_c is proportional to the square of the tangential velocity and is inversely proportional to the radius of rotation. Therefore, for a fixed velocity, the smaller the radius, the larger is the centripetal acceleration and for a fixed radius, the larger the object's velocity, the larger is the centripetal acceleration. The dependence of a_c on the square of v rather than just v makes the influence of v on a_c rather critical. This is an extremely important feature that deserves more attention when various kinds of forces are treated in the later chapters.

3.5 LINEAR VELOCITY, ANGULAR VELOCITY, PERIOD, AND FREQUENCY

An important quantity in this kind of motion is its period. The object's duration, T, for one rotation of its sustained uniform circular motion being constant is called the period of motion. This is determined by dividing the circle's circumference by the object linear velocity. That is,

$$T = \frac{2\pi r}{v}, \tag{3.12}$$

from which

$$v = \frac{2\pi r}{T}. \tag{3.13}$$

The quantity $2\pi/T$ has a special significance that will be discussed in detail in Chapters 8 and 9. It is called the angular velocity of the object. It is denoted by ω and its units are radians per second (rad/s). That is

$$\omega = \frac{2\pi}{T}. \tag{3.14}$$

The inverse of the period, that is, $1/T$, is called the frequency, f, of the motion, which describes how many rotations the object completes in 1 s.

Thus,

$$f = \frac{1}{T}. \tag{3.15}$$

Now, combining Equations 3.13 and 3.14 gives

$$v = \omega r, \tag{3.16}$$

where r is the radius of the orbited circle.

From Equations 3.11 and 3.16, the centripetal acceleration in the uniform circular motion becomes

$$a_c = \omega^2 r. \tag{3.17}$$

EXAMPLE 3.6

Calculate the centripetal acceleration of a racing car driven at a speed of 55.0 m/s (124 mi/h) on a flat circular racing track of 221 m radius.

Solution

From Equation 3.11, we have

$$a_c = \frac{v^2}{r}.$$

Upon substituting for v and r, we get

$$a_c = \frac{(55.0\ \text{m/s})^2}{221\,\text{m}} = 13.7\ \text{m/s}^2.$$

Analysis

Comparing the value of the centripetal acceleration obtained above with the gravitational acceleration g (9.80 m/s²), it turns out that in this example, $a_c = 1.41$ g. Although this is a significant amount, it is still within the limits (maximum of ~5 g) that a human can endure in highly accelerated vehicles. Pilots in the air force and at air shows are frequently subjected to centripetal accelerations much higher than 1.41 g.

EXAMPLE 3.7

A grinding wheel of 0.180 m radius is rotating at an angular speed of 28.0 revolutions per minute. Determine

a. The tangential velocity v of a point on the outer rim of the wheel
b. The centripetal acceleration of a point on the outer rim of the wheel

Solution

a. To find the speed of the rim of the disk, use the circumference and divide that by the time of each revolution. The circumference, C, is

$$C = 2\pi r.$$

Thus,

$$C = 2\pi(0.180\ \text{m}) = 1.13\ \text{m}.$$

The time of each revolution, T, is

$$T = 60.0\ \text{s}/28.0\ \text{revolutions} = 2.14\ \text{s per revolution}.$$

The above value for T is the time that each point on the rim of the disk takes to complete one revolution. Thus, the speed, v, of any point on the rim is given by

$$v = C/T = 1.13 \text{ m}/2.14 \text{ s} = 0.530 \text{ m/s}.$$

b. From Equation 3.11, the centripetal acceleration, a_c, is

$$a_c = \frac{v^2}{r}.$$

Thus,

$$a_c = \frac{(0.530 \text{ m/s})^2}{0.180 \text{ m}} = 1.56 \text{ m/s}^2.$$

Analysis

If the radius of this wheel was half the given value ($r = 0.090$ m), the acceleration of its rim parts would be twice the above value, that is, $a_c = 3.12 \text{ m/s}^2$. Such a high acceleration would require that a grinding disk rotating with the given speed should be made from a highly pressed grinding composite that could withstand against chipping off the rim as the disk is rotated.

EXAMPLE 3.8

An airplane in an air show dives along a portion of a vertical circle of 550.0 m radius with a speed of 140.0 m/s (315.0 mi/h). What is the centripetal acceleration, which the pilot is subjected to during his dive?

Solution

Again, from Equation 3.11, $a_c = v^2/r$, and, upon substituting for v and r

$$a_c = \frac{(140.0 \text{ m/s})^2}{550.0 \text{ m}} = 35.64 \text{ m/s}^2.$$

Analysis

Compare the value of this centripetal acceleration with the gravitational acceleration g (9.80 m/s^2) and notice that $a_c = 3.64$ g. This is a much higher value than the centripetal acceleration experienced by the car in the previous example and is closer to the limit (about 5 g). Of course, decreasing the radius of the dive and/or increasing his velocity would get the pilot much closer to, or higher than, the 5 g alarming limit!

3.6 NONUNIFORM CIRCULAR MOTION

A more involved circular motion is one that has an object rotating in a circular motion, but with a varying velocity, that is, a velocity that is changing in magnitude and in direction. In addition to the centripetal acceleration resulting from the change in the direction of the velocity, the object will experience an acceleration resulting from the change in the magnitude of its tangential velocity. This additional acceleration is called tangential acceleration, denoted by a_t. The tangential acceleration, at any moment of the object's motion, is perpendicular to the centripetal acceleration.

The magnitude of the tangential acceleration, assumed to be constant, is

$$a_t = \frac{\Delta v_t}{\Delta t}.$$

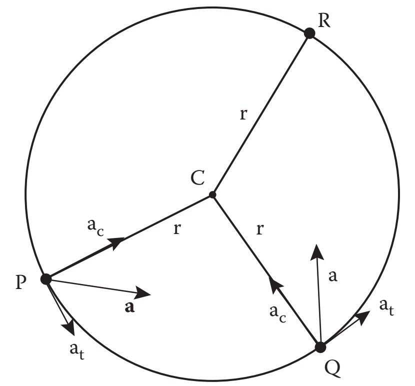


FIGURE 3.7 An object undergoing a nonuniform circular motion experiences at any point two accelerations, a centripetal acceleration a_c toward the center and a tangential acceleration a_t tangent to the circle at that point.

In such situations, the object will have an acceleration with two components, a centripetal acceleration a_c and a tangential acceleration a_t. The total acceleration of the object (Figure 3.7) would then be

$$a = \sqrt{a_c^2 + a_t^2}, \tag{3.18}$$

which will point toward a point off the center by an amount that depends on how large or small the ratio between the centripetal and the tangential acceleration is.

As an example of the nonuniform circular motion, consider Example 3.6 again, but with a slight modification.

EXAMPLE 3.9

A racing car is being driven on a flat circular racing track of 221 m radius with a speed of 55.0 m/s (124 mi/h). However, at a certain point, the racer in a moment's reaction to avoid colliding with the car just ahead of him slowed down his car to 44.0 m/s in an interval of 2.30 s.

a. Calculate the centripetal acceleration of the car at the moment the driver was about to slow down.
b. Calculate the tangential acceleration of the car as the driver was slowing down.
c. Calculate the total acceleration of the car while slowing down.
d. Draw a reasonably clear sketch that would reflect the situation.

Solution

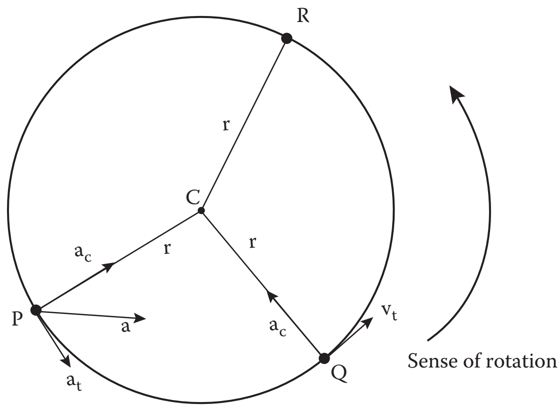
In this case, the car at the location under consideration experiences two accelerations, a centripetal acceleration a_c and a tangential acceleration a_t (the following figure).

a. For the centripetal acceleration, $a_c = v^2/r$, which upon substituting for v and r at P gives

$$a_c = \frac{v^2}{r} = \frac{(55.0\ \text{m/s})^2}{221\,\text{m}} = 13.7\ \text{m/s}^2.$$

This agrees with the result in Example 3.6.

b. For the tangential acceleration, $a_t = \Delta v_t/\Delta_t$,



which upon substituting for $\Delta v_t = (44.0 - 55.0)$ m/s = −11.0 m/s and $\Delta t = 2.30$ s gives

$$a_t = \frac{-11.0 \text{ m/s}}{2.30 \text{ s}} = -4.78 \text{ m/s}^2.$$

c. The magnitude of the total acceleration, a, is

$$a = \sqrt{a_c^2 + a_t^2} = \sqrt{(13.7 \text{ m/s}^2)^2 + (-4.78 \text{ m/s}^2)^2} = 14.5 \text{ m/s}^2.$$

d. The diagram in the figure above demonstrates at point P all parameters relevant to the situation.

Analysis

1. The minus sign of a_t is for the direction of the tangential acceleration being against the motion due to slamming on the brakes.
2. Notice the direction of the total acceleration of this car at point P while the driver was slowing down.

PROBLEMS

3.1 A 12-month baby, Lyla, was observed crawling on a straight line from a location at point A whose coordinates are (0.0 m, 1.0 m) toward a location at point B whose coordinates are (3.2 m, 5.0 m). If the motion between the two locations took 28 s and the baby continued moving at a constant acceleration, determine
   a. The displacement of the baby along the x and y axes
   b. The average velocity of the baby along the x and y axes

3.2 In the previous problem, assuming that the baby's acceleration during her motion was constant, find
   a. The components of her instantaneous velocities along the x and y axes at t = 28 s
   b. The components of her average acceleration during this motion

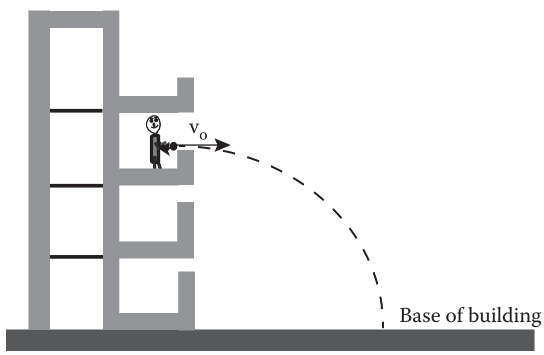
3.3 Consider a bee that has moved on a straight line from position A to another position B such that the x and y components of its velocity at these points were (1.00 m/s, −6.00 m/s) and (1.50 m/s, −12.0 m/s), respectively. Assuming the motion from A to B took 2.00 s, determine
   a. The x-component of the bee's average acceleration

b. The y-component of the bee's average acceleration
c. The magnitude and direction of the bee's average acceleration

3.4 A projectile was fired with an initial velocity of 30.0 m/s in a direction that makes an angle of 30.0° with the horizontal. If air resistance is assumed negligible
a. Find the time it takes the projectile to reach its maximum height and time in air.
b. What is the maximum height in the projectile's trajectory?
c. What is the horizontal velocity of the projectile as it hits the ground?
d. What is the horizontal distance the projectile will travel before it hits the ground?

3.5 A small plastic ball was fired in air from a toy gun with an initial velocity of 5.0 m/s at an angle of 53° with the horizontal. Ignoring air resistance, determine
a. The time it takes the ball to reach its maximum height
b. The x and y components of the ball's velocity after 0.2 s of its motion
c. The x and y coordinates of the ball after 0.2 s of its motion
d. The x and y coordinates of the ball's highest position in its motion

3.6 A boy on a balcony of a seashore building throws a ball horizontally with a speed of 5.0 m/s (the figure below). If the balcony is 9.6 m above the ground, find
a. The time of flight of the ball
b. How far away from the base of the building does the ball land



3.7 Reconsider the figure above in Problem 3.6 and determine
a. The vertical velocity of the ball as it hits the floor
b. The horizontal velocity of the ball as it hits the floor
c. The direction of the velocity of the ball as it hits the floor

3.8 Reconsider Problem 3.6 assuming that as the ball was thrown horizontally, another ball was released from rest. Ignoring air resistance, determine
a. Whether or not the two balls get to the ground at the same time
b. The magnitude and direction of the velocity of each ball as it hits the ground

3.9 In a volleyball game, a player throws the ball from a height of 1.40 m at an angle of 60.0° with a velocity of 8.00 m/s. Determine
a. The velocity of the ball at the top of its trajectory
b. How far from the player does the ball land
c. The magnitude and direction of the velocity of the ball as it hits the ground

3.10 A small rock, secured to a string, is whirled on a vertical circle of radius 35 cm at a constant speed of 1.6 m/s.
a. Find the centripetal acceleration of the rock during its motion.
b. Assuming the string broke when it was at its lowest point, 1.20 m above the ground, find how long it takes the stone to hit the ground and how far from its initial position it hits the ground.

3.11 The average velocity of the electron around the proton in the hydrogen atom is 2.18×10^6 m/s. Knowing that the radius of the orbit in the hydrogen atom is 5.30×10^{-11} m, calculate

a. The period of the electron's motion
b. The electron's centripetal acceleration around the nucleus

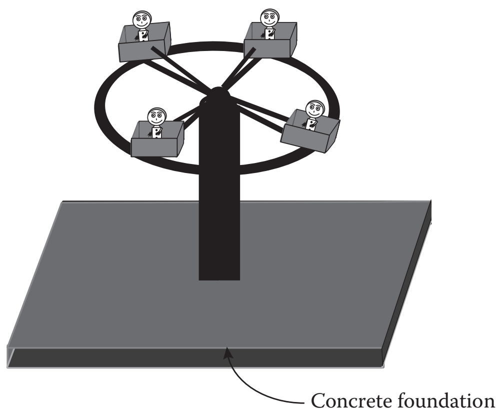
3.12 A small rock of mass 0.15 kg is attached to a string of radius r = 1.10 m. A boy firmly holding on the string at its other end sets the stone into circular motion with a fixed speed such that the stone makes three revolutions every 2.00 s. Determine the radial acceleration of the stone during its motion.

3.13 A satellite is launched to a height of 180 km above the surface of the Earth (Earth's radius is 6380 km) and is set in its orbit at that height with a velocity of 3100 m/s.

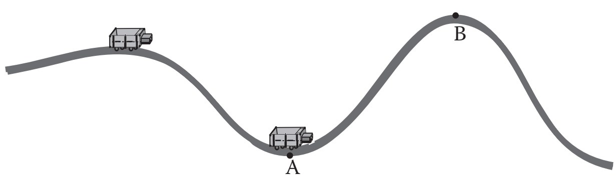
a. Find the time it takes the satellite to complete one revolution around the Earth.
b. Determine the centripetal, that is, the radial acceleration of the satellite in this orbit.

3.14 Each of the passengers in a circular carnival ride in an amusement park (the figure below) circulates at 4.0 m radius from the center of the supporting base. If each passenger makes one complete circle (one revolution) in 8.0 s, calculate

a. The linear, that is, the tangential velocity of any passenger
b. The centripetal acceleration of each passenger
c. The angular velocity, ω, in radians per second and revolutions per minute of any passenger (the figure below).



3.15 A driver is traveling in his sport-utility vehicle (SUV) along a hilly area. He sets the cruise control to a fixed speed of 22 m/s (49.5 mi/h). The road is curved at locations A and B and is of 55 and 45 m radii, respectively. Determine the values and directions of the centripetal acceleration of the car at A and B (the figure below).



4 Newton's Laws
Implications and Applications

So far two kinds of motion have been presented, motion in one and two dimensions. These were described in Chapters 2 and 3. The two chapters dealt with the kinematics of motion. The kinematics of an object results from the forces acting on it. Once a force acts on an object, it creates an acceleration of the object. This acceleration changes the object's velocity in value, or in direction, or both. If such changes are limited to just the magnitude of the velocity, the object then changes its position in a nonlinear form (see Equation 2.12). However, if such changes are limited to just a uniform change in the direction of the velocity, the object then changes its position, tracing a circular path in what is called uniform circular motion (see Section 3.4). In this chapter, the dynamics of objects in a variety of situations will be discussed by dealing with the laws that govern the forces acting on these objects. The treatment comes through three laws stated by Newton that are known as Newton's laws of motion and will be of objects in both linear and circular motions.

4.1 STATEMENT OF NEWTON'S LAWS

Newton formulated three laws of motion that all were named after him in recognition of his insight and distinguished versions of the statements of the laws, especially discovering and formulating the third law that was previously unknown. The following are statements of the Newton's three laws.

4.1.1 First Law

An object, at rest or in motion with a constant velocity, retains that state of being in absolute rest or uniform motion along a straight line unless acted upon by a force.

4.1.2 Second Law

The net force, $\mathbf{F}_{net}$, acting on an object is equal to the time rate of change of its linear momentum, $\mathbf{P}$. A simpler special case version of this law states that the net force acting on an object is equal to the product of the object's mass and its acceleration in the direction of the acting force.

4.1.3 Third Law

When two objects act on each other, they act with forces of equal magnitude and opposite in direction to each other. The two forces are labeled as forces of action and reaction with no distinction between them. Note that these two forces act on different objects.

4.2 DISCUSSION OF NEWTON'S LAWS OF MOTION

4.2.1 Newton's First Law

The substance of the first law can be presented through a broader context that Newton's second law entails. A close look into the second law shows that if there is no net force acting on an object, then the object's linear momentum stays constant. Since momentum, $\mathbf{P} = m\mathbf{v}$, then a constant $\mathbf{P}$ means constant velocity, $\mathbf{v}$. Accordingly, an object that is at rest will remain at rest, and an object moving with a velocity of some initial value and direction would continue its motion with that constant velocity. Therefore, it is clear that the essence of the first law is already contained in the second law, and there is no new information that can be obtained from the first law that would not be obtained from the use of the second law.

4.2.2 Newton's Second Law

The proper mathematical statement of Newton's second law takes the following form:

$$\mathbf{F}_{net} = \frac{\Delta \mathbf{p}}{\Delta t}, \tag{4.1}$$

where $\mathbf{p}$ is the momentum of the object defined as

$$\mathbf{p} = m\mathbf{v}, \tag{4.2}$$

m is its mass. From the above two equations one may have

$$\mathbf{F}_{net} = \frac{\Delta (m\mathbf{v})}{\Delta t}.$$

Assuming that the mass stays constant, this equation reduces to

$$\mathbf{F}_{net} = \frac{m\Delta \mathbf{v}}{\Delta t},$$

or

$$\mathbf{F} = m\mathbf{a}. \tag{4.3}$$

This is a vector equation equivalent in two dimensions to two-component equations:

$$(F_x)_{net} = ma_x, \tag{4.4a}$$

$$(F_y)_{net} = ma_y. \tag{4.4b}$$

EXAMPLE 4.1

Calculate the horizontal force with which a 6-year-old child is pulling her 1.40-kg toy box on a smooth surface so that it moves in a straight line with an acceleration of 1.20 m/s^2.

Solution

Since the motion is one dimensional, say x, the acceleration is then along the x direction. Thus, Newton's second law,

$$(F_x)_{net} = ma_x,$$

which after substitution becomes

$$(F_x)_{net} = (1.40 \text{ kg})(1.20 \text{ m/s}^2) = 1.68 \text{ N}.$$

Analysis

Such a force is moderately small with which, as could be envisioned, a 6-year-old child can move such mass.

EXAMPLE 4.2

How much force is needed to give a truck of 7.00×10^3 kg an acceleration of 5.00 m/s^2?

Solution

Using Newton's second law,

$$(F_x)_{net} = ma_x,$$

the force would then be

$$(F_x)_{net} = (7.00 \times 10^3 \text{ kg})(5.00 \text{ m/s}^2) = 3.50 \times 104 \text{ N}.$$

Analysis

This is a large force that is needed to move a 7.00-ton truck.

4.2.3 Newton's Third Law

Newton's third law is a fundamental law that describes an essential feature of nature. It describes mutual physical forces between any pair of entities when one of them acts with a force on the other. The object that is being acted on by a force responds, that is, reacts, instantaneously with a force of reaction equal and opposite in direction to the force acting on it. That is why these two forces are called action and reaction, and since the two forces act simultaneously, there is no way to distinctly label one of the forces as action and the other as reaction.

One simple example of forces of action and reaction is that between a small book, of mass m, resting on a table (Figure 4.1). The book acts on the table by a force equal to its weight, mg. The table reacts instantaneously and exerts on the book a force of reaction, F_N, equal to mg, and as long as the book is in equilibrium with the table, the two forces mg and F_N are equal. F_N is the normal component of reaction, that is, F_N is the force of reaction normal to the surface of contact, and in this chapter, it is denoted by the symbol F_N.

4.3 FREE BODY DIAGRAM

When applying Newton's second law to an object, one must consider only those forces that act on the object. An example of this can be demonstrated for the book sitting on the table depicted in Figure 4.1. If one is interested in analyzing the dynamics of the book, then the forces acting on only the book are to be considered. In this example, these are the weight of the book acting downward and the reaction on the book exerted upward by the table.

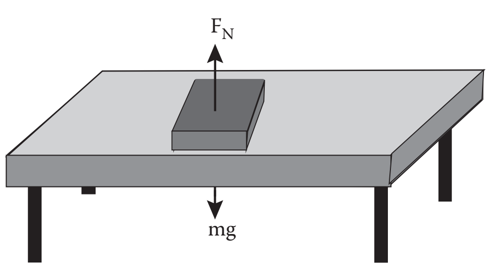


FIGURE 4.1 Illustration of forces of action and reaction acting on a book of mass m resting on a table.

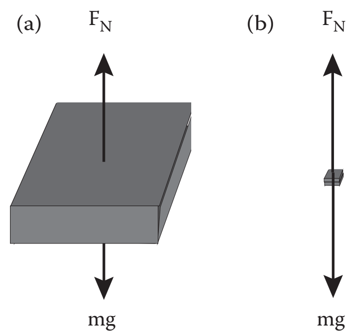
To facilitate simple visualization of this issue, it is advisable to isolate the object from the rest of the system and construct a diagram on which all forces acting on the object are displayed. This method is known as the free body diagram (FBD) for the object.

EXAMPLE 4.3

Draw an FBD for the book in Figure 4.1.

Solution

Notice that only the weight of the book and the normal force of reaction are acting on the book (the figure below). Thus, visualizing the book as a point particle, F_N and mg are the only forces to be displayed in the FBD for the book. This is demonstrated in the figure below (part (b)).



EXAMPLE 4.4

A car of 1.20×10^3 kg was moving with a velocity of 22.0 m/s on a flat straight highway when the driver slammed on the brakes to bring the car to a full stop in 5.00 s.

a. Calculate the average acceleration of the car during the 5.00-s interval.
b. Find the average force acting on the car during the 5.0-s interval.
c. Draw an FBD for the car during the 5.00-s interval.

Solution

a. Using the equation of motion (Equation 2.9)

$$v = v_o + a\,t$$

yields

$$0 = 22.0 \text{ m/s} + a\ (5.00 \text{ s}).$$

Solving for a gives

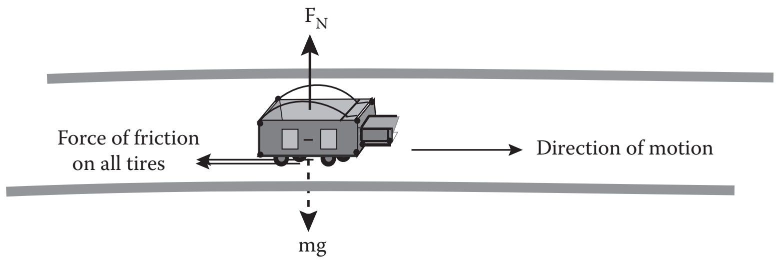
$$a = -4.40 \text{ m/s}^2.$$

Again, one has to remember that the negative sign means the acceleration is against the direction of motion. Thus, it is a deceleration, slowing the car toward its final stop. If we assume that the car's direction of motion is along the positive x-axis, then the car's acceleration, a, points along the negative x-axis.

b. For calculating the net force acting on the car, we have Newton's second law:

$$(F_x)_{net} = ma_x = (1.20 \times 10^3 \text{ kg})\ (-4.40 \text{ m/s}^2) = -5280 \text{ N}.$$

c. The FBD of the car is depicted in the following figure.



Analysis

Notice that the acceleration, being in the direction of the net force, is opposite to the direction of motion.

4.4 REMARKS ON THE APPLICATION OF NEWTON'S SECOND AND THIRD LAWS

1. The acceleration is always in the direction of the net force. Therefore, if only two forces are acting on an object, and they happen to be in opposite directions, then the acceleration is in the direction of the larger force.
2. In cases when several forces act on an object, it becomes necessary to resolve these forces into their components. If these forces are in one plane, as is the case in most of the problems, any two perpendicular directions can be chosen to be the x and y axes. Accordingly, all forces acting on the object should be resolved into their x and y components. Newton's second law can then be applied to the x and y components independently (see Equations 4.4a and 4.4b).
3. Whenever an object is treated as a point particle, it is useful to display all forces acting on the object as arrows with their tails at the object and tips pointing away from the object along the direction of the force. As will be learned in later chapters, this idea does not hold for an extended object known as rigid body.

4.5 NEWTON'S GRAVITATIONAL LAW

Newton's gravitational law demonstrates another inclusive embracement of Newton's third law. According to this law, masses attract each other. For two objects of masses m_1 and m_2 (Figure 4.2), there is a force of attraction $\mathbf{F}_2$ on m_2 exerted by m_1 to which m_2 reacts exerting on m_1 a force $\mathbf{F}_1$, equal but opposite in direction to $\mathbf{F}_2$. And if there are no other forces acting on either of them, the two masses will start moving toward each other under equal forces, $\mathbf{F}_2 = \mathbf{F}_1$, that is, $m_1a_1 = m_2a_2$.

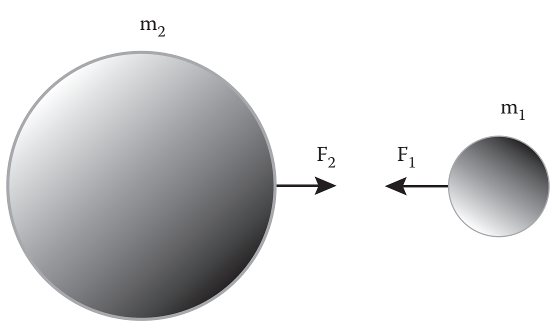


FIGURE 4.2 Demonstration of Newton's third law via depicting equal and opposite forces of attraction and reaction, that two masses m_1 and m_2 exert on each other.

Since these masses are different, their accelerations will be different. The heavier mass will move more slowly than the lighter mass (see Newton's second law).

According to Newton's gravitational law, the force of attraction F between two masses is (1) directly proportional to the product of the two masses and (2) is inversely proportional to the square of the separation between them. A proportionality may be written as an equation by inserting a multiplicative proportionality constant. The constant of proportionality is called the gravitational constant G. Applying this law to an object of mass m in the vicinity of the Earth, mass M_E (Figure 4.3), the force on mass m takes the following form:

$$F_G = G\frac{m\,M_E}{r^2},$$

where r is the distance between the center of mass of the object and the center of the Earth; the gravitational constant $G = 6.67 \times 10^{-11}$ N m²/kg².

As the force on any object according to Newton's second law is F_{net} = mass × acceleration, in the absence of any other forces on mass m, the above force is

$$ma = G\frac{mM_E}{r^2},$$

which leads to

$$a = G\frac{M_E}{r^2}.$$

Using the values for $G = 6.67 \times 10^{-11}$ N m²/kg², $M_E = 5.97 \times 10^{24}$ kg and r, which for objects near the Earth's surface is basically equal to the Earth's radius, $R_E = 6.38 \times 10^6$ m, the acceleration a turns out to be

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

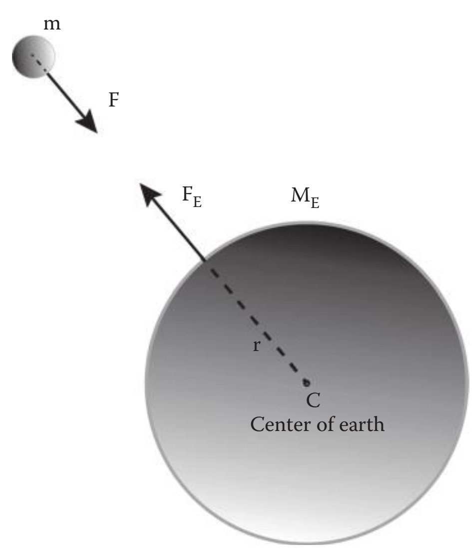


FIGURE 4.3 Demonstration of Newton's gravitation law, applied to a mass m and the Earth, mass M_E.

The above value is that of the gravitational acceleration which any object acquires in a free fall motion.

4.6 MASS, WEIGHT, AND NEWTON'S GRAVITATIONAL LAW

The concept of weight can be understood in the context of Newton's second law. Every object has a mass, which to some extent tells how much material the object has, and mass is a scalar quantity, measured in kilograms when the mks system is used. Therefore, the mass of an object is a fixed quantity that has the same value measured anywhere, on the surface of the Earth or on any other planet. However, the notion "weight" of the object is the force by which the object is attracted to the Earth. This weight can vary rather slightly from one place to another, depending on the object's position on the Earth, because the force of gravity of Earth on objects varies slightly from one location to another. This is so because the force per unit mass on any object, also known as the gravitational acceleration, g, varies slightly from one location to another. The value 9.80 N/kg, that is, 9.80 m/s^2, is adopted as a standard value on or near the Earth, and an object's weight, W, is always equal to the mass of the object multiplied by the gravitational acceleration, g. Thus,

$$W = mg. \tag{4.5}$$

Every planet has its own characteristic gravitational acceleration. On the Moon, it is about 0.17 g, and on Mars, it is 0.38 g. The units of weight are those of any force, that is, Newton, N, and

$$1.0 \text{ N} = 1.0 \text{ kg m/s}^2.$$

EXAMPLE 4.5

An astronaut is of 70.0 kg mass. What is his weight on Earth and what would his weight be on the Moon?

Solution

From Equation 4.5, the astronaut's weight on Earth is

$$W = mg_E = (70.0 \text{ kg})(9.80 \text{ m/s}^2) = 686 \text{ N}.$$

And his weight on the Moon would be

$$W = mg_M = (70.0 \text{ kg})(0.170 \times 9.80 \text{ m/s}^2) = 117 \text{ N}.$$

Analysis

One can notice that although the astronaut's mass is obviously the same, his weight changes significantly on the Moon. The astronaut as he walks on the Moon would feel much lighter than what he would on Earth, simply because the Earth's pull on him is much larger.

Exercise

In a test of the influence of gravitational acceleration on motion, consider an astronaut who, prior to his launch, throws a small stone up from the ground with an initial velocity of 14.7 m/s. Find how high the stone would go in 2.00 s and compare that with the stone's height if it is to be thrown with the same initial speed from the Moon's surface? (see Equation 2.12).

4.7 NEWTON'S THIRD LAW AND APPARENT WEIGHT

A direct application of Newton's third law is the reading of one's weight on a scale (Figure 4.4a). For a boy stepping on the scale, two forces act on him, his weight acting downward and the scale's

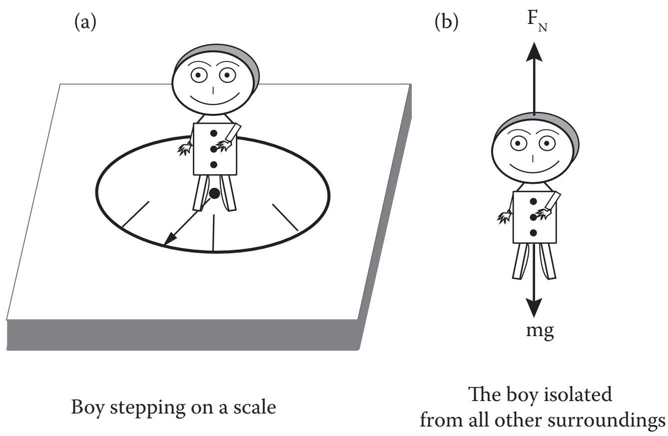


FIGURE 4.4 (a) A boy stepping on a scale and (b) forces acting on the boy, his weight mg downward and the force of reaction F_N upward.

force of reaction, F_N, acting upward (Figure 4.5b). The needle's deflection on the scale represents the force of reaction, which gives the boy's weight.

An interesting aspect of this application occurs when the scale is in an upward or downward accelerated motion. As discussed in the following, the weight that a passenger in such a case would measure is influenced by the elevator's acceleration. The passenger would then find that his or her weight as read on the scale is higher or lower than what it is at rest.

A further comprehension of this effect may be attained by considering the following two cases.

4.7.1 Acceleration Upward

Consider the case of a scale in an elevator accelerated upward with an acceleration a (Figure 4.5a). Applying Newton's second law to the boy, one then has

$$F_{net} = ma.$$

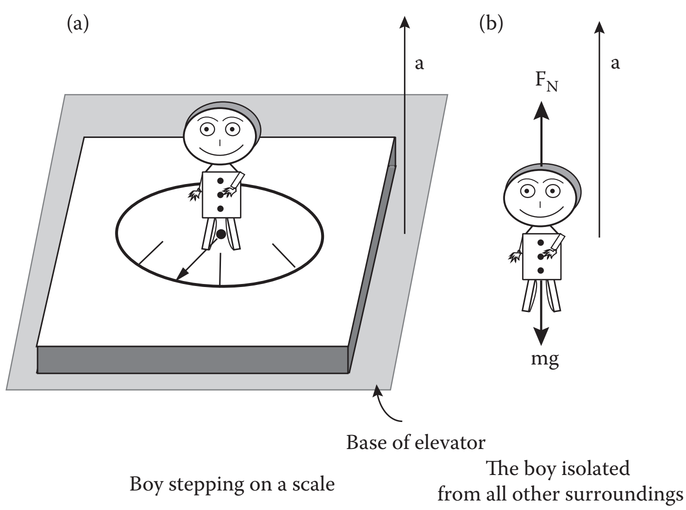


FIGURE 4.5 (a) A scale in an elevator accelerated upward and (b) forces acting on the boy, his weight mg downward and the force of reaction F_N upward.

The net force on the boy is the resultant of his weight acting downward and the force of reaction acting upward (Figure 4.4b). Since the elevator and the boy are moving upward with an acceleration a, then the force of reaction is larger than the weight (see Remarks, Section 4.4). The above equation then becomes

$$F_N - mg = ma.$$

Slight rearrangement gives

$$F_N = mg + ma$$

or

$$F_N = m(g + a). \tag{4.6}$$

That is, the normal force of reaction is larger than the weight mg by a term equal to ma. Since the normal force of reaction is in essence the reading on the scale, this reading is the apparent weight the boy would read on the scale, and hence he would seem heavier than he usually is.

4.7.2 Acceleration Downward

Considering an elevator accelerating downward with an acceleration, a (Figure 4.6a), the net force on the boy would be the resultant of the weight of the boy downward and the force of reaction on the boy upward (Figure 4.6b). Since the elevator, and the boy, are moving downward with an acceleration a, the boy's weight is larger than the force of reaction on him (see Remarks, Section 4.4). Thus, Newton's second law becomes

$$F_{net} = ma$$

and then becomes

$$mg - F_N = ma.$$

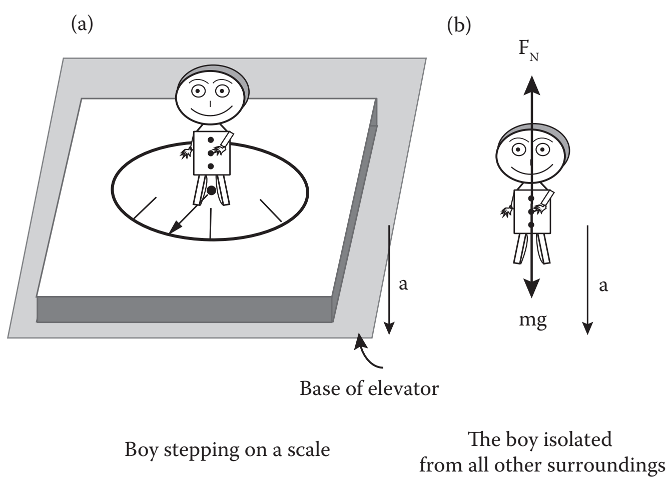


FIGURE 4.6 (a) A scale in an elevator accelerated downward and (b) forces acting on the boy, his weight mg downward and the force of reaction F_N upward.

After slight rearrangement, the above equation becomes

$$F_N = mg - ma$$

or

$$F_N = m(g - a). \tag{4.7}$$

That is, the normal force of reaction is smaller than the normal weight mg by a term equal to ma. Again since the normal force of reaction is in essence the reading on the scale, this reading is the weight, better called the apparent weight, the boy would read on the scale, and hence he would seem lighter than he usually is.

Special Case: Notice that in this case as the elevator accelerates with a larger, a, the smaller the boy's apparent weight becomes less, and in the limit of $a = g$, the apparent weight would be zero. That is, the boy becomes "weightless" as read on the scale!

EXAMPLE 4.6

A girl of 66.0 kg mass, on her way to a physical fitness facility in the third floor of a complex, used the building elevator on her way up and on her way down after doing her usual exercises. There is a scale in the elevator that one may use. Find the girl's apparent weight on her way up and down if in both cases the elevator had an acceleration of 3.00 m/s².

Solution

On the way up, use Equation 4.6. Thus,

$$F_N = m\,(g + a) = (66.0\text{ kg})(9.80\text{ m/s}^2 + 3.00\text{ m/s}^2) = 845\text{ N}.$$

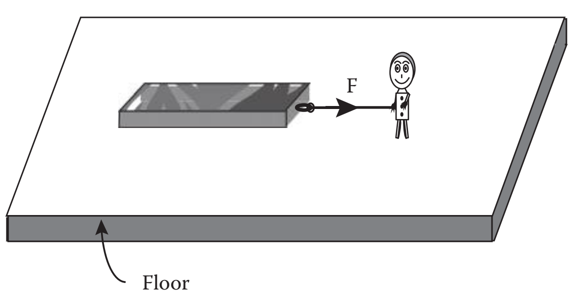
On the way down, use Equation 4.7. Thus,

$$F_N = m\,(g - a) = (66.0\text{ kg})(9.80\text{ m/s}^2 - 3.00\text{ m/s}^2) = 449\text{ N}.$$

EXAMPLE 4.7

A 2.00-kg toy, initially at rest (the figure below), is being pulled across a smooth level porch with a constant force of 2.70 N by a child via a string attached to the toy, parallel to the floor.

a. Find the acceleration of the toy as it moves on the porch.
b. Find the normal component of the force of reaction of the ground on the toy. State its direction.
c. Determine the magnitude and direction of the force acting on the child's hand.
d. Determine how far the toy moves in 3.00 s.



Solution

a. From Newton's second law applied in the x direction (the figure below, part (a)),

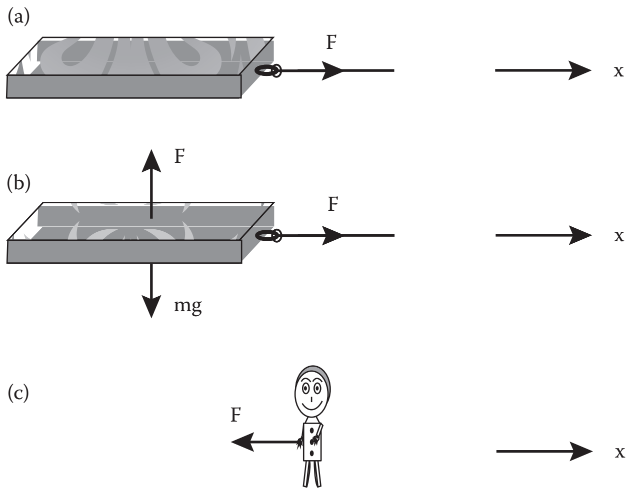
$$F_{net} = ma_x.$$

Thus,

$$2.70 \text{ N} = (2.00 \text{ kg})a.$$

This gives

$$a_x = 1.35 \text{ m/s}^2.$$



b. The toy is not accelerated upward nor downward. In fact, the toy is not moving at all in the vertical direction (the figure above, part (b)). Thus, using Newton's second law in the vertical direction,

$$F_{net} = ma_y,$$

one then has

$$F_N - mg = 0,$$

$$F_N = mg = (2.00 \text{ kg})(9.8 \text{ m/s}^2) = 19.6 \text{ N}.$$

c. The force of reaction on the child's hand is equal to the force with which the child is pulling on the toy, that is, 2.70 N (the figure above, part (c)). This force is exerted on the child's hand by the toy and is in the negative x direction.
d. For this part, one needs to resort to the equation of motion (Equation 2.12):

$$x = v_0 t + \left(\frac{1}{2}\right)at^2.$$

This gives

$$x = (0)(3.00 \text{ s}) + \left(\frac{1}{2}\right)(1.35 \text{ m/s}^2)(3.00 \text{ s})^2 = 6.08 \text{ m}.$$

Analysis

Note here that it is the force on the toy, getting it accelerated, and hence moving through a certain displacement in the 3.00-s interval.

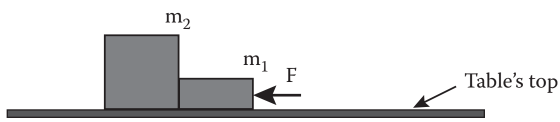
Exercise

As a further test on the blend between the dynamics of an object and its kinematics, test yourself by finding the velocity of the toy at the end of the 3.00-s interval. (*Hint*: See Equation 2.9.)

EXAMPLE 4.8

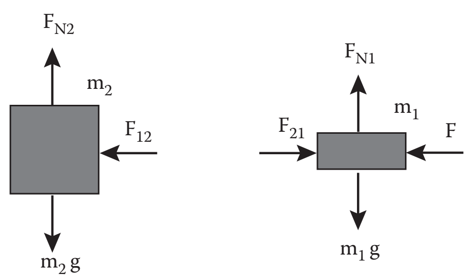
Two blocks of masses $m_1 = 4.20$ kg and $m_2 = 8.40$ kg are set on a smooth table's top surface next to each other (the figure below). Mass m_1 was pushed horizontally by a constant force $F = 25.2$ N, after which the two blocks started moving as one system.

a. Draw an FBD for each of the blocks.
b. Determine the acceleration of each block.
c. Determine the force of contact between the two blocks.

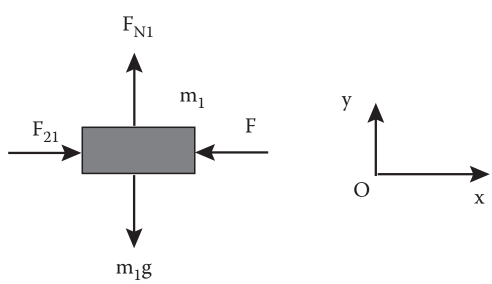


Solution

a. The FBD diagram for each of the blocks is depicted in the figure below. All forces acting on each block are sketched as stated in the statement of the exercise.



b. Visualizing block 1 as a point particle with an x–y coordinate system whose origin is at the point particle block, and applying Newton's second law along the x direction, gives (the figure below).



For the first block,

$$F_{net,x} = m_1a.$$

That is,

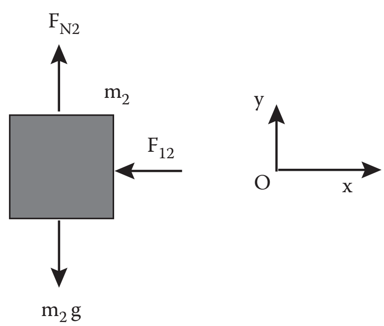
$$F - F_{21} = m_1 a.$$

Substituting for all given quantities gives

$$25.2\ \text{N} - F_{21} = (4.20\ \text{kg})a. \tag{4.8}$$

For the second block (the figure below),

$$F_{net,x} = m_2 a.$$



That is,

$$F_{12} = m_2 a.$$

Substituting for all given quantities gives

$$F_{12} = (8.40\ \text{kg})a. \tag{4.9}$$

Adding Equations 4.8 and 4.9 gives

$$25.2\ \text{N} - F_{21} + F_{12} = (4.20\ \text{kg})a + (8.40\ \text{kg})a.$$

Since F_{12} and F_{21} are equal in magnitude, they add to zero, and

$$25.20\ \text{N} = (12.60\ \text{kg})a.$$

Thus,

$$a = 2.00\ \text{m/s}^2.$$

This is the acceleration of each of the blocks. It has to be the same for both blocks because the two are moving as one system and not as two separate entities.

c. Substituting for a in Equation 4.9 gives

$$F_{12} = (8.40\ \text{kg})\ (2.00\ \text{m/s}^2) = 16.8\ \text{N}.$$

Accordingly, $F_{21} = 16.8$ N. It is just pointing in a direction opposite to that of F_{12}.

Analysis

Here are several points that deserve attention:

1. The pushing force F is shown acting on block m_1 only. That does not mean that it has no influence on block m_2. In fact, its influence is clear in creating a contact force (force of action) F_{12} on block 2. If the pushing force F was absent, the contact force F_{12} would not be present either.
2. The force of action F_{12} on block 2 generates a force of reaction F_{21} on block 1.
3. Each of the two forces of action and reaction, as sketched in the two separate FBD diagrams, acts on only one block: F_{21} on block 1 and F_{12} on block 2, and hence each is an important part of the dynamics of each block. Notice that the second subscript of each of the forces is used here to indicate the block on which the force is acting, while the first subscript labels the source of the force.
4. In the treatment of the dynamics along the x-axis, there was no mention of the dynamics along the y-axis, simply because the problem asked for the acceleration of each block along the x-axis, and that is completely independent, in this case, of the dynamics of the blocks along y-axis.

General Remarks

As was stated in one of the special remarks earlier, the forces of action and reaction between two constituents of a system such as the system treated in the above exercise are both important when these two blocks are treated separately. However, if the system is handled as a whole, then the forces of action and reaction are considered internal forces that do not influence the dynamics of the system. In other words, when the system is not split into its separate parts, these internal forces cancel each other.

Exercise

For the system described in the previous example, determine the force of reaction the table's top surface exerts on each block. (*Hint*: as neither of the blocks is accelerating in the y direction, the acceleration of each block along the y-axis is zero. Hence, applying Newton's second law to each block gives F_N for each block.)

EXAMPLE 4.9

A block of mass m is released to slide down a smooth incline of length L, height $h = 8.00$ m, and angle $\theta = 30°$ above the horizontal. Find the velocity of the block at the bottom of the incline, point B.

Solution

As the motion is along the incline, it would be convenient to have the x-axis aligned along the incline pointing toward B (the following figure). The only external force acting on the block is its weight pointing vertically downward. This is resolved into two components $F_x = mg\sin\theta$ and $F_y = -mg\cos\theta$. There is the normal force of reaction from the incline on the block along the y-axis.

From Newton's second law applied in the x direction

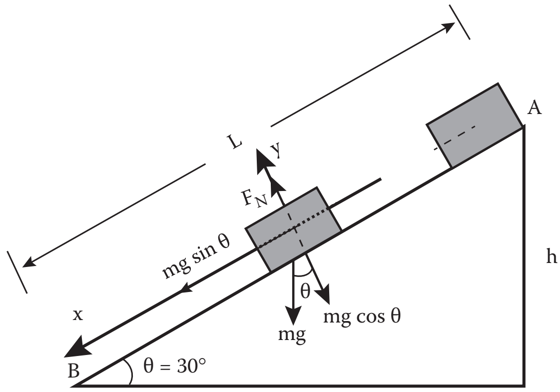
$$F_{net} = ma_x.$$

Thus,

$$mg\sin 30.0° = ma_x.$$

This gives

$$a_x = (g)\sin 30.0° = (9.8\ \mathrm{m/s^2})(0.50) = 4.9\ \mathrm{m/s^2}.$$



For finding the velocity, the equation of motion, $v^2 = v_o^2 + 2\,ax$, after substituting for the known quantities gives

$$v^2 = 0 + 2\,g\sin\theta\,L = v^2 = 2\,g\left(\frac{h}{L}\right)L = 2\,gh.$$

Thus, the block's velocity at B is

$$v = \sqrt{2gh} = \sqrt{2\,(9.80\,m/s^2)(8.0\,m)} = 13\,m/s^2.$$

4.8 DYNAMICS OF A UNIFORM CIRCULAR MOTION

In both uniform and nonuniform circular motions discussed in Chapter 3, it was demonstrated that the centripetal acceleration, also called radial acceleration, of an object results from the change in the direction of its velocity. As forces are the cause of motions (Newton's second law), there has to be a net centripetal force, acting on the object, and hence creating the centripetal acceleration. It is that force that enables a driver to negotiate a ramp and constantly change the direction of the car's velocity until he or she is off the ramp. In situations when roads are icy, the acting "centripetal" forces are significantly small, and obviously negotiating a ramp with the commonly familiar speeds of 30 or 40 mi/h becomes very difficult or sometimes impossible. Treatment of frictional forces between a moving object and the surface on which the object is moving, a car on a curved ramp included, will be considered in detail in the next chapter. The current discussion will be limited to cases where frictionless surfaces are involved. It is the net centripetal force, acting on the object, that sustains the object's circular motion.

The sum of the radial forces determines the car's centripetal acceleration,

$$a_c = \frac{v^2}{r}.$$

From Newton's second law, the radial force is

$$F_{net,c} = ma_c = m\frac{v^2}{r}.$$

The right-hand side of the above equation gives the magnitude of this centripetal force. But, what is its origin? It originates from the source of this force. In fact, there are numerous interesting

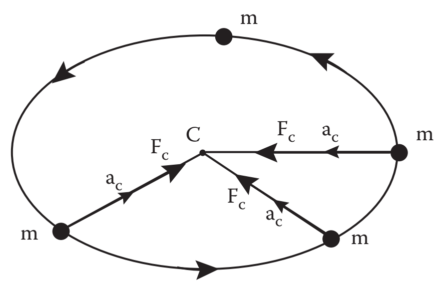


FIGURE 4.7 An object of mass m circulating in a uniform horizontal circular motion. The net centripetal force F_c acting on the object is the centripetal force.

situations that demonstrate the origin and role of these forces. Figure 4.7 is a typical diagram that shows an object of mass m circulating in uniform horizontal circular motion. The centripetal force F_c in the diagram is the net force acting on the object and keeps it circulating at a fixed distance r from the center of the circle.

EXAMPLE 4.10

A small block of mass m = 321 g is attached firmly through a locked hook to a string of length 0.450 m. The mass then is set into a horizontal circular motion by a boy whirling the block at a constant speed of 1.20 m/s.

a. Determine the centripetal acceleration of the circulating mass.
b. Determine the centripetal force acting on the circulating mass.
c. What is the source of the centripetal force acting on the circulating mass?

Solution

a. From Equation 3.11 the centripetal acceleration is

$$a_c = \frac{v^2}{r}.$$

Thus,

$$a_c = \frac{(1.20\,\text{m/s})^2}{0.450\,\text{m}} = 3.20\,\text{m/s}^2.$$

b. The centripetal force is

$$F_{net,c} = ma_c = (0.321\text{ kg})(3.20\text{ m/s}^2) = 1.03\text{ N}.$$

c. The centripetal force in this situation is the tension, which the boy is exerting on the string as he is whirling it.

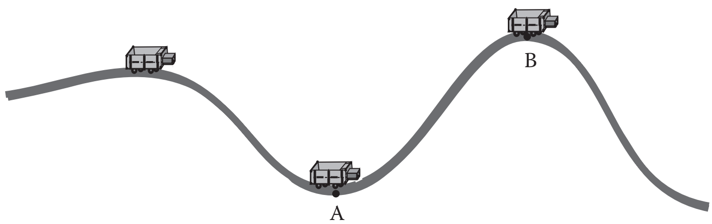
Analysis

1. The above value for the tension on the mass is reasonably feasible for the boy to exert. If he intends to go to much higher speeds, he has to apply more force through his grip on the string.

2. From Newton's third law, the mass will react on the boy's hand and acts on it with a force equal to the tension the boy is exerting on the mass via the string. That is why the boy feels strained after a while, and in case he wants to go to much higher speeds, he has to exert more force, and accordingly bears more burden; otherwise, he simply has to give up this ambition.

EXAMPLE 4.11

A driver, in his SUV, is travelling in a hilly area. He sets the cruise control to a fixed speed of 18.0 m/s (40.5 mi/h). The road is curved in several places among which are locations A and B (the figure below). The van is 1.80×10^3 kg, and the curvatures at the two locations are 60.0 and 50.0 m, respectively.



a. Sketch an FBD diagram of the vertical forces on the van at locations A and B.
b. Determine the values and directions of the centripetal acceleration on the van at A and B.
c. Determine the reaction of the road on the van at both locations. (*Hint*: For this part, refer to the dashed circles sketched at both locations so that you can tell in what direction the centripetal acceleration at each location should be pointing.)

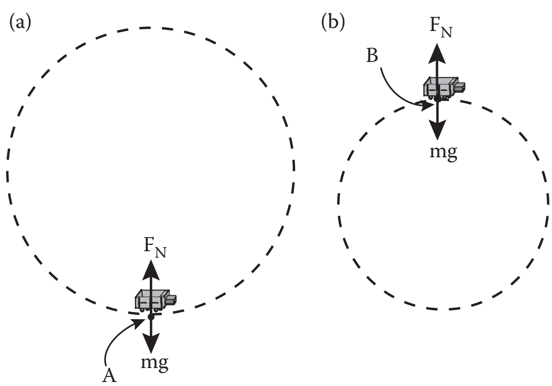
Solution

a. The FBDs of the car at A and B are shown in the figure below (part (a) and (b)), respectively.
b. The accelerations of the van at points A and B are as follows:

At A:

$$a_c = \frac{(18.0\ \text{m/s})^2}{60.0\ \text{m}} = 5.40\ \text{m/s}^2$$

directed toward the center of the circle passing through A, thus directed upward.



At B:

$$a_c = \frac{(18.0\,\text{m/s})^2}{50.0\,\text{m}} = 6.48\,\text{m/s}^2$$

directed toward the center of the circle passing through B, thus directed downward.

c. As the centripetal acceleration at location, A, is toward the center, that is, upward, the force of reaction F_N is larger than the weight mg. Thus, Newton's second law,

$$F_{net,c} = ma_c,$$

applied at A becomes

$$F_c = m\frac{v^2}{r}.$$

That is,

$$F_N - mg = m\frac{v^2}{r},$$

From which,

$$\begin{aligned} F_N &= mg + m\frac{v^2}{r} = (1.80\times10^3\,\text{kg})(9.80\,\text{m/s}^2) + (1.80\times10^3\,\text{kg})(5.40\,\text{m/s}^2) \\ &= 1.76\times10^4\text{N} + 0.972\times10^4\text{N} = 2.73\times10^4\text{N}. \end{aligned} \tag{4.10}$$

However, the centripetal acceleration at B is downward because the center of the circle at location B is below the sketched curvature (see the figure above, part (b)), which means that the weight mg is larger than the force of reaction F_N. Thus, Newton's second law,

$$F_{net,c} = m\,a_c,$$

applied at B becomes

$$F_c = m\frac{v^2}{r}.$$

That is,

$$mg - F_N = m\frac{v^2}{r},$$

from which

$$\begin{aligned} F_N &= mg - m\frac{v^2}{r} = (1.80\times10^3\,\text{kg})(9.80\,\text{m/s}^2) - (1.80\times10^3\,\text{kg})(6.48\,\text{m/s}^2) \\ &= 1.76\times10^4\text{N} - 1.17\times10^4\text{N} = 0.590\times10^4\text{N}. \end{aligned} \tag{4.11}$$

Analysis

1. The force of reaction on the car at A is greater than its value at B. That is due to the direction of the curvature being upward at A and downward at B.
2. Notice that F_N at B (Equation 4.11) is much less than F_N at A (Equation 4.10). If the velocity was increased a little bit from its given value, the force of reaction F_N at B could be zero or even negative. Zero value for F_N means the car, the passengers included cease to feel their weight any more. The car is then barely touching the ground. Once the value of F_N at B becomes negative, the car is lifted above the ground and becomes airborne.

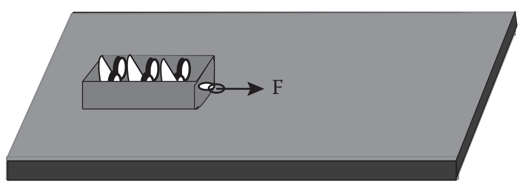
Exercise

Try the above example for another value of the velocity of the SUV being 22.1 m/s (79.6 mi/h) at both A and B. Check if your results correlate with the second comment of the above analysis.

PROBLEMS

4.1 A 0.80-kg toy (the figure below), initially at rest, is being pulled across a smooth level floor by a child who exerts a constant force of 4.0 N via a string directed parallel to the floor.

a. Find the acceleration of the toy as it moves on the floor.
b. Determine how fast the toy will be moving after 3.0 s.
c. Find the normal component of the force of reaction of the floor on the toy. State its direction.



4.2 A block of mass m = 7.00 kg is pulled horizontally on a flat frictionless table's top by a constant force F (= 14.0 N) as shown in the figure below.

a. Determine the acceleration of the block.
b. Determine the normal force of reaction that the table exerts on the block.
c. Draw an FBD for this block.

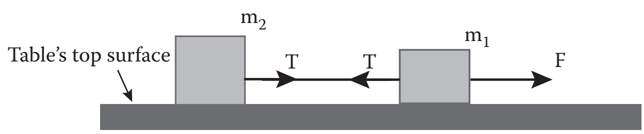
4.3 A 110 N box of wheat is being pulled across a level floor at a constant speed by a force of 30.0 N directed 37.0° above the horizontal (the figure below).

a. Draw an FBD of the wheat box.
b. Calculate the normal component of the reaction of the ground on the box.
c. Determine the frictional force between the box and the ground's surface.

4.4 A child accidentally bumped into a 1.40-kg wooden toy box that got set into motion along a straight line on the surface of a tiled floor. The box was slowed down at an average rate of 0.900 m/s^2. Determine the force of friction acting on the box during its motion.

4.5 Two blocks of masses $m_1 = 4.20$ kg and $m_2 = 8.40$ kg are set on a smooth table's top surface and are attached to each other by a string, as shown in the figure below. If the first block is pulled horizontally to the right with a force $F = 34.0$ N,

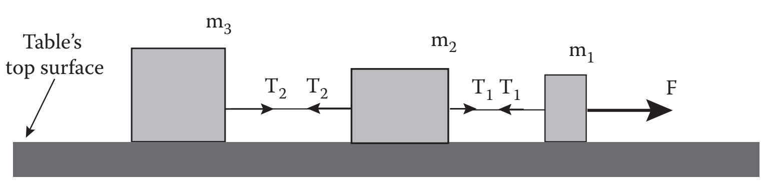
a. Determine the acceleration of each block.
b. Determine the tension in all connecting strings.



4.6 Reconsider the previous problem, treating the two blocks as one system.

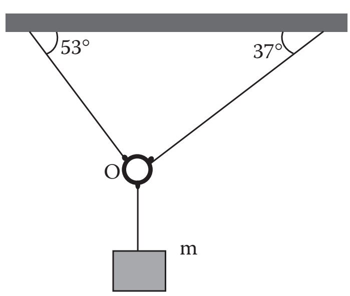
4.7 Three blocks of masses $m_1 = 4.20$ kg, $m_2 = 8.40$ kg, and $m_3 = 12.6$ kg are set on a smooth surface and are attached to each other by two strings, string 1, connecting blocks 1 and 2, and the other, string 2, connecting blocks 2 and 3 as shown in the figure below. If one pulled horizontally on the first block to the right with a force $F = 34.0$ N,

a. Determine the acceleration of each block.
b. Determine the tension in all connecting strings.

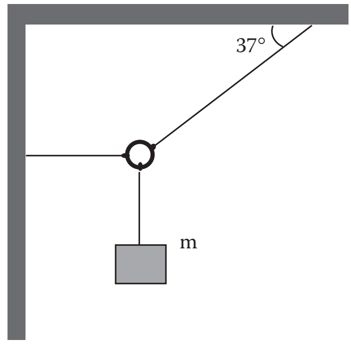


4.8 Reconsider the previous problem treating the three blocks as one system.

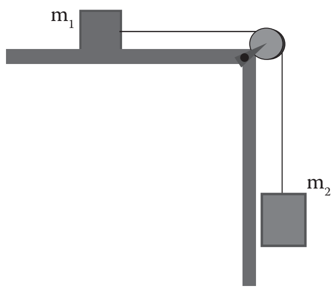
4.9 In the figure below, two strings are attached to a ceiling support and a mass $m = 5.50$ kg is attached to a third string such that the three strings are tied to a light metal ring O. The entire system is in equilibrium. Determine the tension in each string.



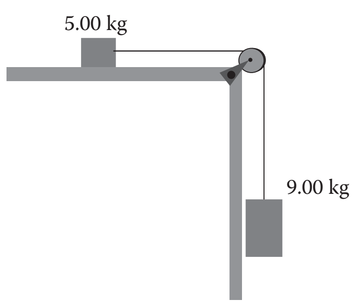
4.10 The following figure shows three strings, one is attached to a ceiling and another to a vertical wall support. A mass $m = 11.0$ kg is hanging vertically downward from the third string such that the strings are all tied to a light metal ring O. The system is in equilibrium. Determine the tension in each string.



4.11 A system of two masses with $m_1 = 3.00$ kg, resting on a frictionless horizontal table, and $m_2 = 6.00$ kg hanging vertically via a string that passes over a frictionless pulley and is attached to m_1 (the figure below).
a. Determine the acceleration of each block.
b. Determine the tension in the cable connecting the two blocks.

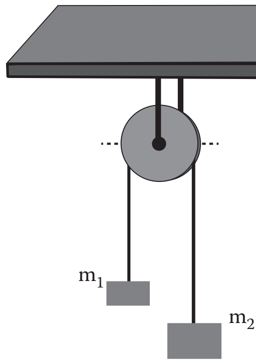


4.12 A 9.00-kg hanging mass is connected via a string passing over a pulley to a 5.00-kg block that slides on a flat frictionless table (the figure below). The pulley is frictionless. Determine
a. The acceleration of the system
b. The tension in the string



4.13 The following figure displays a well-known textbook problem called the Atwood machine. In this case, two masses $m_1 = 4.00$ kg and $m_2 = 8.00$ kg are connected via a string that passes over a frictionless pulley. The pulley is attached to the ceiling through a central axis as part of a U-shaped solid support.
a. Determine the acceleration of the two masses.
b. Find the tension in the string connecting the two masses.

c. Assuming that the string is long enough, find the velocity of mass m_2 after it has traveled 0.25 m below its initial position.



4.14 A 30.0-kg child rides on a Ferris wheel that rotates in a vertical plane with a speed $v = 6.00$ m/s. The Ferris wheel has a radius of 8.0 m. Determine

a. The centripetal acceleration that the child experiences during this ride

b. The centripetal force acting on the child during this ride

4.15 The electron (mass, $m = 9.11 \times 10^{-31}$ kg) in a hydrogen atom rotates in a circular orbit of a radius r that is the closest to the nucleus with a velocity of 2.20×10^6 m/s.

a. What would be its centripetal acceleration if the orbit's radius $r = 5.29 \times 10^{-11}$ m?

b. What would be the centripetal force acting on the electron during its motion?

5 Newton's Laws
Friction in Linear and Circular Motions

Newton's laws have been introduced in Chapter 4, where they were applied to numerous cases of motion. Naturally, there are many kinds of forces that generate motion. They are elastic forces such as tension and spring forces, resistive forces such as air resistance and frictional forces that impede motion, gravitational forces such as those manifested by weights of objects, and collision forces that are common examples of the forces usually treated in a mechanics textbook. As Newton's laws are universal for velocities much smaller than the speed of light, they can be applied to all kinds of forces, including resistive forces such as friction. In this chapter, friction is handled in detail in two main kinds of motion, linear motion and circular motion. The dynamics of linear and circular motions where friction is involved displays several interesting features that warrant a separate treatment. This is discussed separately in this chapter instead of including it in Chapter 4.

5.1 FRICTIONAL FORCES

This chapter details frictional forces involved in numerous linear and circular motions. In linear motion, the frictional force is directed opposite to the direction of motion or to the intended direction of motion. However, in an arbitrary circular motion, the frictional force can have two directions, one opposing the motion and another perpendicular to the circular path, that is, directed toward the center of the object's circular path. The former type causes the object to slow down along the path of motion, while the second type of frictional force is due to the rough nature of the two contacting surfaces, and in essence, it supplies the centripetal force necessary for sustaining the object's motion along a circular path. Both of these frictional forces are addressed in the following sections.

5.1.1 Frictional Forces: Linear Motion

When a block of mass, say m, is laid on the surface of a table, there exists a potential for frictional forces to appear between the two surfaces as the block is pulled or pushed across the table. If one intends to pull on the block on the table's surface via a force T, an opposition is felt due to a relative roughness of the two surfaces. Of course, the smoother the surfaces of contact the easier it becomes to move the block on the table's top.

It is a common observation that as one increases the pulling force applied to move the block, the block still holds onto its position without movement, up to a point. The frictional force described in this case is called static force of friction and is denoted by F_s. This situation is represented in Figure 5.1.

However, as T_1 continues to increase, there comes an instant when the pulling force reaches a certain value at which the block starts to move. This indicates that during the no response status, the opposing frictional force, acting on the block, grows in magnitude in a manner that negates the effect of the applied pulling force up to a limit at which the frictional force is exceeded by the applied force and the block then starts to move. The frictional force described in this case is called kinetic force of friction and is denoted by F_k. In addition, observations confirm that once the block starts moving, the pulling force, T_1, could be slightly decreased to a smaller value T_2, and one could still move the block along the table's surface (Figure 5.2).

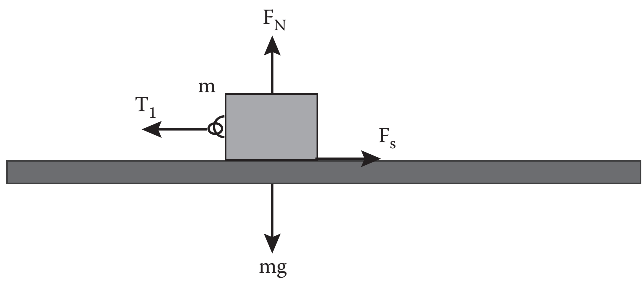


FIGURE 5.1 A block on which a pulling force T_1 is acting to the left, prior to the start of its motion. In this case, $T_1 < F_s$.

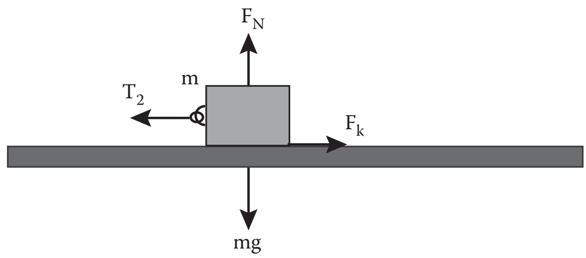


FIGURE 5.2 A block after it starts moving when the pulling force $T_2 \geq F_s$.

The above observations imply that there are two kinds of frictional forces: a static frictional force, F_s, which acts on the block prior to motion, and a kinetic force of friction, F_k, which acts on the object after it starts moving, and $F_s > F_k$. In each case, the force of friction is proportional to the normal force, F_N pushing the surfaces together. F_N is always normal to the surfaces of contact between the two contacting objects. The constant of proportionality between the frictional force and the normal force of reaction is called the coefficient of friction and is labeled as μ_s or μ_k, depending on whether the case is static or kinetic, respectively. Therefore, for the two cases, one can write the following two relations:

a. Static force of friction (Figure 5.1):

$$F_s = \mu_s F_N. \tag{5.1}$$

b. Kinetic force of friction (Figure 5.2):

$$F_k = \mu_k F_N. \tag{5.2}$$

And, as described earlier,

$$F_k < F_s. \tag{5.3}$$

That is,

$$\mu_k < \mu_s. \tag{5.4}$$

Note that the normal component of the force of reaction between two objects may differ from one case to another, depending on the orientation of their surfaces. For example, in Figure 5.3, the normal force of reaction F_N between m and the horizontal surface is

$$F_N = mg. \tag{5.5}$$

However, if the same block is placed on a surface that is oriented on an incline (Figure 5.4), the normal force of reaction between m and the inclined surface is

$$F_N = mg \cos \theta,$$

which is less than mg.

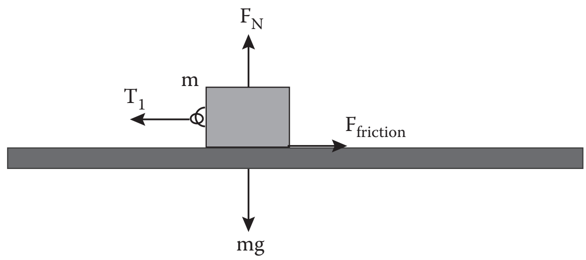


FIGURE 5.3 An object in motion on a flat surface under the action of a horizontal force. The normal force of reaction is equal to the object's weight.

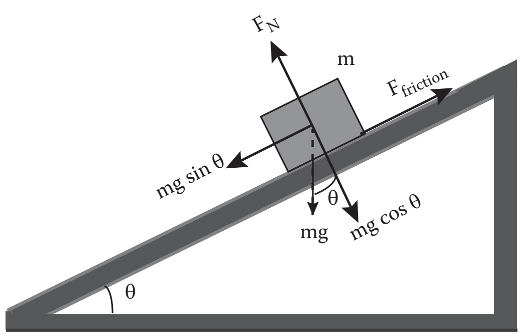


FIGURE 5.4 An object on an incline in motion under gravity. The normal force of reaction is less than the object's weight.

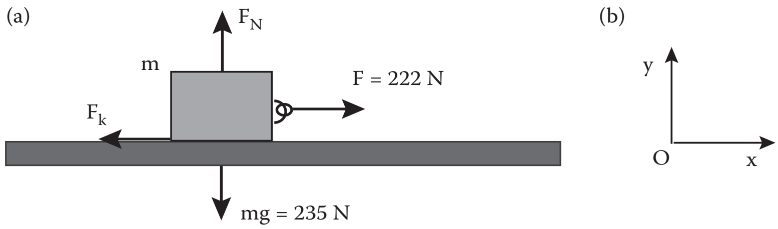
EXAMPLE 5.1

A bookstore worker is pushing horizontally with a force of 222 N a large box over a flat rough surface. Knowing that the mass of the box is 24.0 kg and the coefficient of kinetic friction between the box and the ground is 0.250, calculate

a. The normal component of the force of reaction of the floor on the box
b. The acceleration with which the box would move

Solution

As assumed earlier, any object under consideration is represented by a point at which all acting forces should be sketched or visualized as arrows emerging from the point-particle object and pointing in the designated direction. Accordingly, the above example would be solved in the same way giving the same answer if the worker was pulling on the box from its right side instead of pushing it from the left (the figure below).



a. The force of kinetic friction is

$$F_k = \mu_k F_N.$$

To find F_N, one may apply Newton's second law to the box (the figure above, part (a)), in the vertical direction, that is, the y direction.

In the vertical direction: $F_{net,y} = ma_y$. Since in the vertical direction, $a_y = 0$, then

$$F_N - mg = 0.$$

That is,

$$F_N - (24.0\ \text{kg})(9.80\ \text{m/s}^2) = 0$$

and

$$F_N = (24.0\ \text{kg})(9.80\ \text{m/s}^2) = 235\ \text{N}.$$

b. Applying Newton's second law to the box in the horizontal direction gives

$$F_{net,x} = ma_x$$

and

$$F_k = \mu_k F_N = (0.250)(235\ \text{N}) = 58.8\ \text{N}.$$

Thus,

$$222\ \text{N} - F_k = (24.0\ \text{kg})a_x.$$

That is,

$$222\ \text{N} - 58.8\ \text{N} = (24.0\ \text{kg})a_x.$$

Solving for a_x gives

$$a_x = 6.80\ \text{m/s}^2.$$

Comment

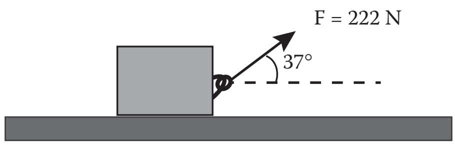
The normal component of the force of reaction F_N that was found in part (a) is the force with which the ground acts on the box. It is a reaction to the force (box weight) with which the box acts on the floor.

In view of the above remark, Example 5.1 can be repeated with a slight modification to the angle along which the box is being pulled.

EXAMPLE 5.2

A bookstore worker is pulling at an angle of 37.0° above the horizontal a large box over a flat rough surface with a force of 222 N (the figure below). Knowing that the mass of the box is 24.0 kg and the coefficient of kinetic friction between the box and the ground is 0.250, calculate

a. The normal component of the force of reaction of the floor on the box
b. The acceleration with which the box would move



Solution

a. Consider the box with all forces acting on it as depicted in the figure below (part (a)). As the force of kinetic friction is

$$F_k = \mu_k F_N,$$

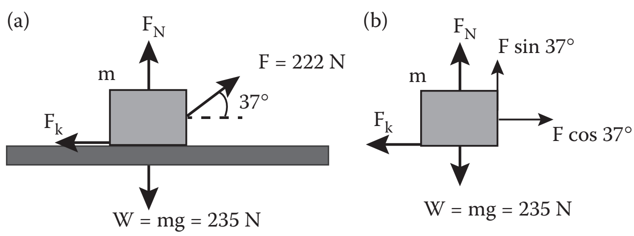
to find F_N, apply Newton's second law to the box in the vertical direction. However, before doing that, define all forces acting on the box along the y direction. These are F_N, the vertical component of the worker's force F in the positive y direction (= F sin θ), and the box's weight in the negative y direction (see FBD, the figure below, part (b)).

For Newton's second law in the vertical direction

$$F_{net,y} = ma_y.$$

Thus,

$$F_N + F \sin\theta - mg = 0, \ a_y = 0.$$



That is,

$$F_N + (222\ \text{N})(\sin 37.0°) - (24.0\ \text{kg})(9.80\ \text{m/s}^2) = 0,\ F_N + 134\ \text{N} - 235\ \text{N} = 0,$$

which after rearrangement gives

$$F_N = 101\ \text{N}.$$

b. To find the acceleration of the box along the x-axis, we need to apply Newton's second law to the box in the horizontal direction. The forces acting on the box along the x-axis are: the horizontal component of the worker's force F (= F cos θ) and the force of friction F_k (= 25.3 N). Thus, Newton's second law in the positive *x* direction becomes

$$F_{net,x} = ma_x.$$

Thus,

$$F \cos \theta - F_k = (24.0 \text{ kg})a_x.$$

Now

$$F_k = \mu_k F_N = (0.250)(101 \text{ N}) = 25.3 \text{ N}.$$

That is,

$$(222 \text{ N}) (\cos 37.0°) - 25.3 \text{ N} = (24.0 \text{ kg})a_x.$$

Solving for a_x gives

$$a_x = 6.33 \text{ m/s}^2.$$

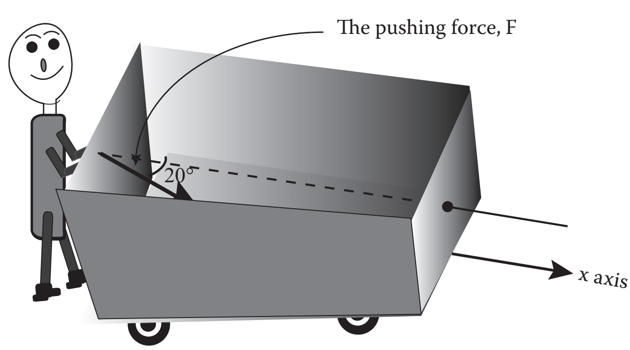
Analysis

1. The normal component of the force of reaction F_N that was found in part (a) of this example is less than its value as determined in the previous example. The pulling force with an angle above the horizontal has a portion, that is, a component, acting upward that lessens the full action of the weight of the box on the ground, and hence the ground reacts with less force than it would if the pulling force were just horizontal.
2. The acceleration of the object along the positive axis in this example is less than its value in the previous example. Does this make sense?

EXAMPLE 5.3

A shopping cart (the figure below) with its load of 5.00 kg combined mass is being pushed across a level slightly rough floor of a super market by a force of F = 9.00 N at a 20° angle below the horizontal. Knowing that the coefficient of friction between the cart and the floor is 0.100,

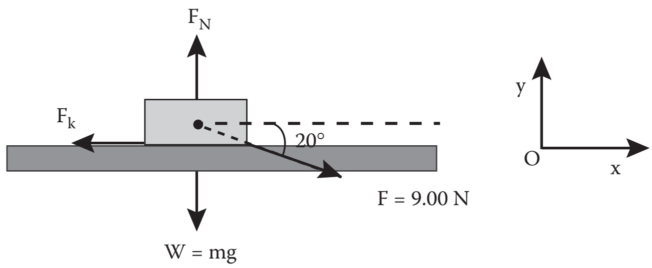
a. Draw an FBD diagram for this cart
b. Calculate the normal component of the reaction of the floor on the cart
c. Determine the frictional force between the cart and the floor's surface
d. Determine the acceleration of the cart



Solution

a. Visualizing the three-dimensional cart as a small rectangular box, the FBD would look as in the figure below.

b. The normal component of the force of reaction, F_N, can be determined from applying Newton's second law in the vertical direction. The forces acting along the y-axis are the normal component of the force of reaction F_N acting upward, the y component of the applied force, $F_y = 9.00 \sin 20°$, pressing down on the floor, and the weight of the cart mg acting downward.



As the cart is not accelerated in the vertical direction, in fact not moving at all in the y direction, then

$$a_y = 0,$$

and

$$F_{net,y} = ma_y$$

becomes

$$F_N - ((mg + (9.00\text{ N})(\sin 20°)) = 0.$$

Thus,

$$F_N = ((5.00\text{ kg})(9.80\text{ m/s}^2) + 3.08\text{ N}) = 52.1\text{ N}.$$

c. The frictional force in the negative *x* direction is

$$F_k = \mu_k F_N.$$

Thus,

$$F_k = (0.100)\ (52.1\text{ N}) = 5.28\text{ N}.$$

d. The acceleration of the cart in the forward, that is, the x direction can be calculated from Newton's second law

$$F_{net,x} = ma_x.$$

Thus,

$$F \cos\theta - F_k = (5.00\text{ kg})a_x.$$

That is,

$$(9.00\text{ N})\cos 20.0 - 5.28\text{ N} = (5.00\text{ kg})a_x.$$

Solving for a_x gives

$$a_x = 0.635\text{ m/s}^2.$$

Analysis

Notice that the force of reaction of the floor on the cart is larger than the cart weight, simply because there is part of, that is, a component of, the pushing force that is pressing down on the floor. The floor in turn is reacting to both, the cart weight and the vertical component of the pushing force combined.

5.1.2 Frictional Forces: Circular Motion

Among the most interesting motions that involve friction are those of vehicles on curved roads and ramps. The interest and importance comes from the fact that, in so many cases, motion itself is not practically possible without the presence of the proper forces toward the center of the curved path along which a vehicle is traveling. In cases where friction is underestimated or possibly ignored, and an over the limit speed is pursued, a skid off the road is most likely. To further one's understanding of this issue, a visit to Sections 3.4 and 4.8 could be very helpful.

In both uniform and nonuniform circular motions discussed in Chapter 3, it was demonstrated that the centripetal acceleration in a circular motion of an object results from the change in the direction of the velocity of the object. As forces are the cause of motions (Newton's second law), there has to be a net centripetal force that is acting on the object, creating the centripetal acceleration. It is that force that lets the driver change the velocity of the car along the curved ramp. In situations when roads are icy, the "proper," that is, "centripetal," forces are significantly smaller, and negotiating a curve with the familiar speeds of 30 or 40 mi/h becomes more difficult if not impossible.

To demonstrate the situation further, consider a car that is moving on a flat circular path (Figure 5.5) with a velocity of constant value (see Section 3.4). From the discussion presented for uniform circular motion, the car experiences a centripetal acceleration, a_c $(= v^2/r)$. Thus, there is a force toward the center that the car should be subjected to, and it is this force that causes the centripetal acceleration. From Newton's second law, this force is equal to

$$F_{net,c} = ma_c = m\left(\frac{v^2}{r}\right). \tag{5.6}$$

The right-hand side of the above equation gives the magnitude of this centripetal force. From where does this acceleration originate? It has to come from a force that is directed toward the center. Such an invisible force is what keeps the car moving along the circular track. That force is a force of friction between the tires and the road. As no single point of the car tires is in constant contact with the road, this friction then is the static friction.

This force is given by

$$F_s = \mu_s F_N.$$

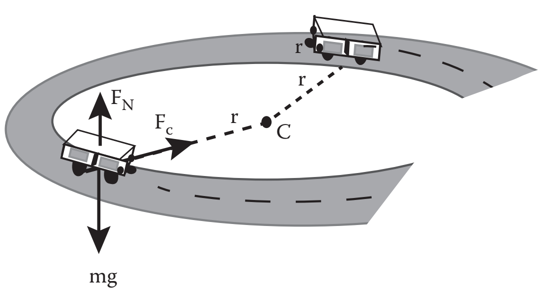


FIGURE 5.5 A car moving on a flat circular track. The centripetal force comes from the static force of friction between the tires and the track.

Hence, Equation 5.6 may be rewritten as

$$\mu_s F_N = m\left(\frac{v^2}{r}\right), \tag{5.7}$$

where μ_s is the coefficient of static friction between the tires and the road.

Applying Newton's second law in the vertical direction ($a_{vertical} = 0$),

$$F_{net,vertical} = ma_{vertical}$$

leads to

$$F_N - mg = 0.$$

Thus,

$$F_N = mg. \tag{5.8}$$

Substituting for F_N from Equation 5.8 into Equation 5.7 gives

$$\mu_s mg = m\left(\frac{v^2}{r}\right),$$

which upon rearrangement becomes

$$v^2 = \mu_s gr.$$

Thus,

$$v = \sqrt{\mu_s gr}. \tag{5.9}$$

Equation 5.9 gives the velocity in terms of the coefficient of friction between the tires and the road, the gravitational acceleration, and the radius of the curve along which the car is moving. Since the gravitational acceleration is constant, the only factors affecting the car's speed would be the radius of the curved road and the coefficient μ_s, which reflects the conditions of the road and the tires. The smaller the radius of the curved road, that is, the sharper the curve, the lower v should be, and the poorer conditions of the road that make μ_s practically small, the worse it is for the driver to negotiate the curve. If the conditions were so bad such that roads were covered with solid ice, the velocity would have to be zero, which means driving on curved roads in extremely icy conditions is simply not possible.

EXAMPLE 5.4

A racing car of 1.20×10^3 kg is driven on a flat circular racing track of 3.00×10^2 m radius. Assuming that the coefficient of static friction between the tires of the car and the racing track is 0.660, calculate

a. The safe upper limit which the driver cannot exceed
b. The centripetal acceleration of the car
c. The centripetal force acting on the car

Solution

a. The limit of the speed as given by Equation 5.9 is $v = \sqrt{m_s gr}$.
   Thus,

$$v = \sqrt{(0.660)\,(9.80\,\text{m/s}^2)\,(3.00 \times 10^2\,\text{m})} = 44.1\,\text{m/s}.$$

b. The centripetal acceleration is given by (see Equation 3.11)

$$a_c = \frac{v^2}{r}.$$

   Upon substituting for v and r,

$$a_C = \frac{(44.1\,\text{m/s})^2}{3.00 \times 10^2\,\text{m}} = 6.48\,\text{m/s}^2.$$

c. The centripetal force acting on the car is

$$F_c = ma_c.$$

   Thus,

$$F_c = (1.20 \times 10^3\text{ kg})(6.48\text{ m/s}^2) = 778\text{ N}.$$

Analysis

Comparing the value of the centripetal acceleration obtained to the gravitational acceleration $g = 9.80$ m/s^2, it turns out that, in this example, $a_c = 0.660$ g. This is a reasonably modest value. Although the coefficient of static friction is relatively high, the radius of curvature has limited the value of the safe speed to only 44.1 m/s (99.1 mi/h).

EXAMPLE 5.5

Consider a driver in a small car of 7.50×10^2 kg who is about to take an exit off a flat sharp ramp posting a sign that reads 30.0 mi/h (13.3 m/s) for a maximum speed on this ramp. The radius of the ramp is 2.00×10^2 m.

a. Determine the minimum coefficient of static friction required between the road and the tires.
b. Calculate the centripetal acceleration that the car experiences as it moves at the maximum possible speed.
c. Calculate the centripetal force that the car experiences as it moves at the maximum possible speed.

Solution

a. Again, the limit of the speed as given by Equation 5.9 is

$$v = \sqrt{\mu_s gr}.$$

   Thus,

$$13.3\,\text{m/s} = \sqrt{\mu_s (9.80\,\text{m/s}^2)\,(2.00 \times 10^2\,\text{m})}.$$

Solving for μ_s gives

$$\mu_s = 0.090.$$

b. The centripetal acceleration is given by (Equation 3.11)

$$a_c = \frac{v^2}{r}.$$

Upon substituting for v and r,

$$a_C = \frac{(13.3 \text{ m/s})^2}{2.00 \times 10^2 \text{ m}} = 0.884 \text{ m/s}^2.$$

c. The centripetal force is given by (Equation 5.6)

$$F_c = ma_c = m\frac{v^2}{r}.$$

Thus,

$$F_c = (7.50 \times 10^2 \text{ kg})(0.884 \text{ m/s}^2) = 663 \text{ N}.$$

Analysis

Comparing the obtained value of the centripetal acceleration with the gravitational acceleration g (9.80 m/s²), it turns out that in this example, $a_c = 0.090$ g, that is, less than one-tenth of the gravitational acceleration. Thus, the centripetal force acting on the car due to friction is unsurprisingly rather small.

5.2 BANKED ROADS

In the last section and the last two examples, it was found that the presence of a centripetal frictional force is mandatory for providing sustainable safe driving on flat-curved roads. The larger the value such a centripetal frictional force has, the safer it becomes to drive at higher speeds along the curve. Thus, any additional force directed toward the center would allow the speed of the circulating vehicle to have even higher values. Such a possibility could come from adding to the ramps some banking such that the road would look higher on the outside edge of the ramp than on the inside of the ramp. The road is then said to be banked. In the following, two situations, (A) banked roads that are frictionless and (B) banked roads that provide some friction, and a discussion of a (hypothetical) frictionless banked roads are presented.

A. Figure 5.6 demonstrates the situation of a car on a banked-curved road. The forces that act on the car while it is on the ramp are the weight of the car vertically downward toward the center of the Earth and the force of reaction, F_N, by the ramp on the car. This reaction force can be resolved into two components, one vertically upward that balances the weight of the car and another component directed horizontally toward the center of the curved ramp. One may label the vertical direction as y and the horizontal direction as the centripetal.

Applying Newton's second law in the two directions, the vertical upward and the horizontal directions, gives the following:

1. The vertical direction:

$$F_{net,\,y} = ma_y.$$

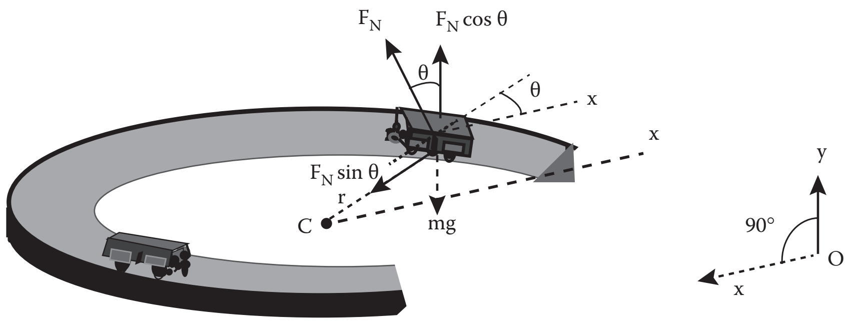


FIGURE 5.6 A car moving on a frictionless banked road experiences a centripetal force that comes totally from the component of the force of reaction, $F_N \sin\theta$.

That is,

$$F_N \cos\theta - mg = 0;\ \ a_y = 0.$$

Thus,

$$F_N \cos\theta = mg. \tag{5.10}$$

2. The centripetal direction:

$$F_{net,c} = ma_c.$$

Since

$$a_c = \frac{v^2}{r},$$

then

$$F_N \sin\theta = m\frac{v^2}{r}. \tag{5.11}$$

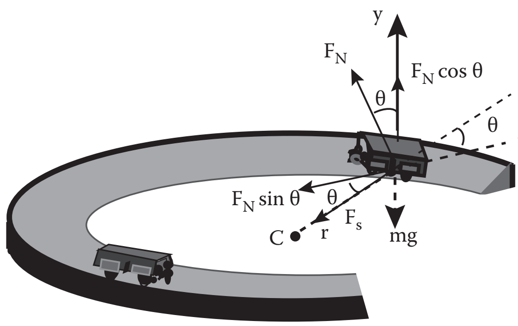
Combining Equations 5.10 and 5.11 gives

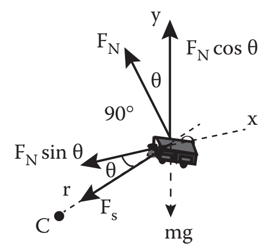
$$\tan\theta = \frac{v^2}{rg}. \tag{5.12}$$

The above equation represents the condition for the banking angle of the road that allows the car to be driven at a speed v on a ramp of radius r and negotiates the ramp safely. As shown in Equation 5.12, the banking angle does not depend on the mass of the car. Therefore, all sizes of vehicles, light or heavy, need the same angle of banking for a safe driving on a frictionless road.

Notice that this argument has not included friction at all. Therefore, with any finite value of friction, one would expect that v could even be higher than what Equation 5.12 allows for. This is elaborated on in the following.

B. The figure below illustrates the banked road, the car, and the forces acting on it. The inset on the right-hand side of the figure is an illustration of axes x and y, x being the horizontal direction and y vertically upward.





The static force of friction is along the banked road, perpendicular to F_N. It makes an angle θ below the horizontal. Thus, the static force of friction F_s along the banked road will assume its maximum value if the speed of the vehicle is high enough to get it very close to skidding his case,

$$F_{s\,max} = \mu_s F_N. \tag{5.13}$$

This can be resolved into two components, a horizontal and a vertical one along mg. The component of F_s along the horizontal direction is

$$F_{sx} = F_s \cos\theta = \mu_s F_N \cos\theta. \tag{5.14a}$$

and along the vertical is

$$F_{sy} = \mu_s F_N \sin\theta. \tag{5.14b}$$

The x and y components of the normal force of reaction are

$$F_{Nx} = F_N \sin\theta, \tag{5.14c}$$

$$F_{Ny} = F_N \cos\theta. \tag{5.14d}$$

The equilibrium in the y direction implies that $F_{ynet} = 0$. That is,

$$F_{Ny} - (mg + F_{sy}) = 0.$$

Substituting for F_{Ny} and F_{sy} from Equations 5.14d and 5.14b in the above equation, it becomes

$$F_N \cos\theta - (mg + \mu_s F_N \sin\theta) = 0,$$

which reduces to

$$F_N = \frac{mg}{(\cos\theta - \mu_s \sin\theta)}. \tag{5.15}$$

The maximum value of the net available horizontal force is then

$$F_{xnet} = F_{Nx} + F_{sx} = F_N \sin\theta + \mu_s F_N \cos\theta. \tag{5.16}$$

It is this value that the upper limit of the centripetal force $m(v^2/r)$ should be bound to, which in this case

$$m\frac{v^2}{r} = F_N \sin\theta + \mu_s F_N \cos\theta.$$

$$m\frac{v^2}{r} = F_N (\sin\theta + \mu_s \cos\theta).$$

Upon using Equation 5.15, this becomes

$$m\frac{v^2}{r} = \frac{mg(\sin\theta + \mu_s \cos\theta)}{(\cos\theta - \mu_s \sin\theta)}.$$

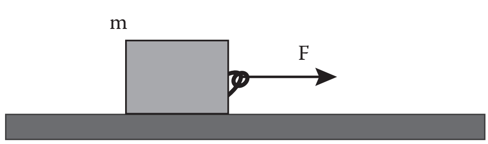
That is, the upper limit of v is

$$v = \sqrt{\frac{gr\,[(\sin\theta + \mu_s \cos\theta)}{(\cos\theta - \mu_s \sin\theta)]}}. \tag{5.17}$$

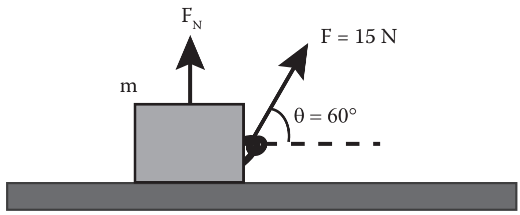
Any value for v below that given by Equation 5.17 would make it safer for the driver to negotiate the ramp.

PROBLEMS

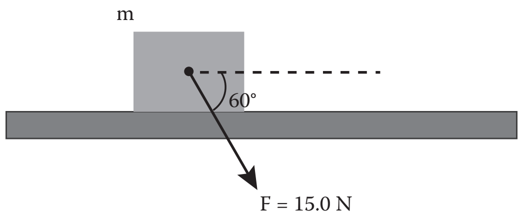
5.1 A string is used to pull a wooden block of a mass m = 2.60 kg across a rough surface with a horizontal force F (the figure below). If the coefficients of static and kinetic friction between the block and the surface are 0.160 and 0.140, respectively, find
  a. The normal force of reaction on the block
  b. The value of the force of friction when the applied force F happens to be equal to one of the following values: 3.0 N, 4.08 N, or 15.0 N
  c. The value of the acceleration when F = 15.0 N



5.2 Letting the force F (= 15.0 N) in the previous problem point upward at an angle of 60.0° above the horizontal (the following figure), and for a coefficient of kinetic friction of 0.140, find
  a. The normal force of reaction and force of friction acting on the block
  b. The acceleration of the block



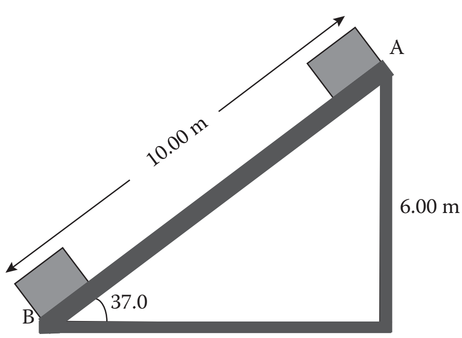
5.3 Consider the previous problem, where the 2.6-kg block is being pushed with a force of 15.0 N in a direction of 60.0° below the horizontal (the figure below). With a kinetic coefficient of friction of 0.100, determine
   a. The normal force of reaction of the floor on the block
   b. The force of friction acting on the block
   c. The acceleration of the block



5.4 In a massive furniture store inventory, a large sofa of 50.0-kg mass is being pushed horizontally at constant velocity across a flat floor with a force of 107.8 N. Determine
   a. The force of friction acting on the sofa
   b. The force of reaction on the sofa
   c. The coefficient of friction between the sofa and the floor

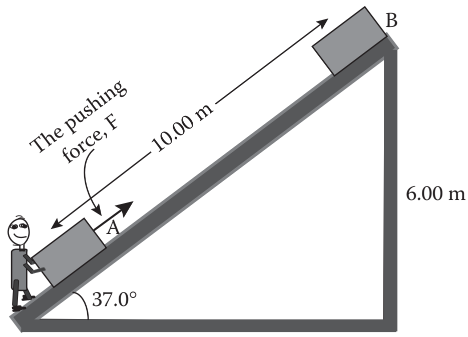
5.5 In a bookstore, a worker pushes a large box over a flat rough surface with a force of 125 N directed 30.0° below the horizontal. Knowing that the mass of the box is 24.0 kg and the coefficient of kinetic friction between the box and the ground is 0.250,
   a. Determine the force of reaction exerted on the box by the yard's surface.
   b. Calculate the acceleration with which the box moves.
   c. Find the distance the box moves in 4.00 s, assuming that the box was initially at rest.

5.6 A 10.0-kg box initially at rest at A slides 10.00 m down toward B on a ramp inclined at an angle of 37.0° with the horizontal (the figure below). The coefficient of kinetic friction between the ramp surface and the box is 0.140.
   a. Draw an FBD for the box.
   b. Determine the acceleration with which the box would move.
   c. Find the time it takes the box to get to position B.
   d. Calculate the velocity of the box at the bottom of the ramp at B.

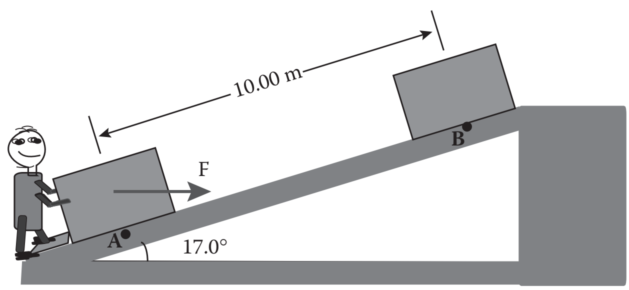


5.7 Reconsider the previous problem in a reverse manner. A 14.0-kg box initially at rest at A is being pushed up the incline by a force directed parallel to it (the figure below), such that the box is moving with constant velocity. The coefficient of kinetic friction between the ramp surface and the box is 0.140,

a. Draw an FBD for the box.
b. Find the force with which the box is pushed.

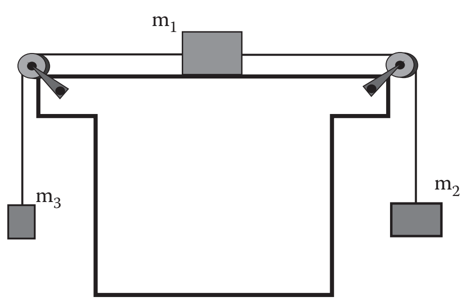


5.8 A factory worker pushes horizontally on a 14.0-kg box to move it up an incline as shown in the figure below. If the coefficient of kinetic friction between the ramp surface and the box is kept unchanged (0.140), find the worker's force F that he should apply so that the box would move with constant velocity.



5.9 An object with mass $m_1 = 3.00$ kg, resting on a smooth horizontal table, is connected to two strings, each of which passes over a pulley on either side of the table's top. Each string is then fastened to a hanging object with mass $m_2 = 8.00$ kg and $m_3 = 4.00$ kg (the figure below). The masses, initially at rest, start to move, with m_2 downward and m_1 to the right.

a. Determine the acceleration of each block.
b. Determine the tension in the strings connecting blocks m_1 to m_2 and m_1 to m_3.



5.10 Redo the previous problem keeping all data the same (the figure above) except for considering a force of friction $F_k = 4.00$ N between m_1 and the table's top. The masses, initially at rest, start to move, m_2 downward, m_1 to the right, and m_3 upward.

a. Determine the acceleration of each block.

b. Determine the tension in the cables connecting blocks m_1, m_2 and m_1, m_3.

5.11 A piece of gum was observed stuck to a disk, 8.00 cm from its center. If the disc is rotating horizontally about a vertical axis through its center at a rate of 45.0 rev/min, find the minimum coefficient of friction that is necessary for the gum to remain on the disk's surface without flying off.

5.12 A racing car goes into a sharp curb of 280 m radius with a speed of 45.0 m/s (101 mi/h).

a. Find the minimum coefficient of friction so that the car experiences a force of friction as high as 0.600 of the car's weight.

b. Would the value for the minimum coefficient obtained above be good enough for the car to negotiate the curb safely?

5.13 What should be the safe speed for a 7250-kg car to negotiate circular road of 111 m radius if the driver has spotted an oily patch on the road, where the coefficient of static friction was only 0.221. Calculate the centripetal acceleration that the car experiences as it moves at the maximum possible safe speed.

5.14 On one of the highways, there was an exit with a flat circular ramp of 2.50×10^2 m radius. A car of 1.20×10^3 kg entered the exit at a velocity of 15.0 m/s. If the driver kept going on the ramp with a constant speed, what should be

a. The minimum coefficient of friction that has to be present for the car to negotiate the ramp safely?

b. The minimum force of friction that would be acting on the car to negotiate the ramp safely?

5.15 Consider a road that is extremely icy and hence is frictionless. If the road is banked at a certain angle θ,

a. What should be the angle of banking of this ramp (radius $r = 110$ m), to enable the driver to drive his small car (mass $m = 7.50 \times 10^2$ kg) with a safe speed of 13.3 m/s (30 mi/h)?

b. Calculate the centripetal acceleration the car experiences as it moves at the maximum possible speed.

c. Calculate the centripetal force that the car experiences as it moves at the maximum possible speed.

6 Work and Energy

As has been established in the previous chapters, Newton's second law is an effective tool for analyzing the dynamics of moving objects. However, to apply Newton's laws to a moving object, all forces acting on the object have to be specified in magnitude and in direction. In many cases, an alternative approach based on energy considerations can be followed. The energy-based approach is effective and very helpful because much of the formalism is cast in terms of scalars, not vectors. This approach starts with a definition of work done on the object by a force or forces acting on it. In addition to the kinetic energy that a moving object has, another energy concept, called potential energy, is introduced. This concept is linked to the work done on the object by a special kind of forces, called conservative forces, among which gravitational forces are a good example. Another quantity, called the total mechanical energy, which is defined as the sum of an object's kinetic and potential energies, is also introduced and discussed in a variety of examples.

6.1 WORK

The work, W, done by a force, F, to displace an object a distance d (Figure 6.1) is defined as

$$W = Fd\cos\varphi, \tag{6.1}$$

where φ is the angle between F and d.

If the force and the displacement are in the same direction (Figure 6.2), then the angle is 0°, and W becomes

$$W = Fd,\ \ \cos 0° = 1.$$

Note that it is only the magnitudes of both the force and displacement d that enter into Equation 6.1; φ is the angle between the directions of the acting force and the resulting displacement. The unit of work is Joule (J).

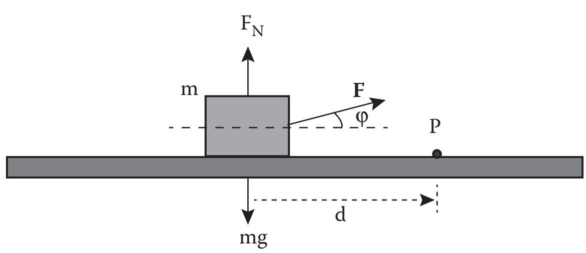


FIGURE 6.1 A block displaced a distance d on a horizontal surface by a constant force **F** directed at angle φ above the horizontal.

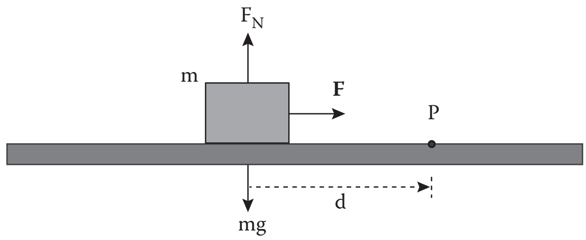


FIGURE 6.2 A block on a horizontal surface acted on by a horizontal force **F**.

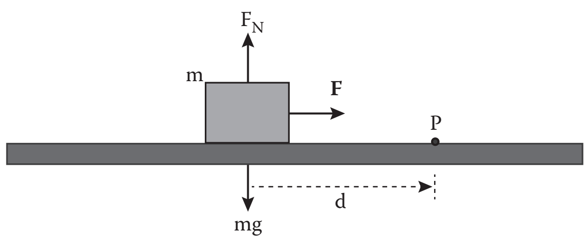
EXAMPLE 6.1

An 8-year-old child is pulling her 2.80-kg toy box on a smooth surface with a force F = 3.60 N via a horizontal string, moving the box in a straight line 4.20 m with an acceleration of 1.20 m/s². Determine

a. The amount of work the child does on the box during this motion
b. The amount of work the weight of the box does during this motion
c. The amount of work the force of reaction does on the box during this motion
d. The total work done on the box during this motion

Solution

A sketch of the box and the forces acting on it is shown in the figure below.



a. For finding the work done by F, use Equation 6.1:

$$W_F = Fd \cos \varphi,$$

where F = 3.60 N, d = 4.20 m, and $\varphi = 0.00°$.
Thus,

$$W_F = (3.60\ \text{N})(4.20\ \text{m})(\cos 0.00°) = 15.1\ \text{J}.$$

b. The work that the weight of the box, mg, does during this motion is

$$W_{mg} = (F_{mg})d \cos \varphi.$$

Thus,

$$W_{mg} = [(2.80\ \text{kg})(9.80\ \text{m/s}^2)](4.20\ \text{m})(\cos 90°) = 0.00\ \text{J}.$$

c. For finding the work that the force of reaction does on the box during this motion, again Equation 6.1 is used:

$$W_{FN} = F_N d \cos \varphi.$$

Thus,

$$W_{FN} = F_N(4.20\ \text{m})(\cos 90°) = 0.00\ \text{J}.$$

d. To find the total work done on the block during this motion, add all contributions of work done by all forces. Thus,

$$W_{total} = W_F + W_{mg} + W_{FN} = 15.1\ J + 0.00\ J + 0.00\ J = 15.1\ J.$$

Analysis

1. It is clear that in this example only the pulling force F worked on the box.
2. Forces that are perpendicular to the displacement of an object do no work on it. This is what was found in parts (b) and (c). Also, in this case, the force of reaction is perpendicular to the displacement, so the value of the force of reaction need not be known. It does no work, since $\cos 90° = 0$

6.2 WORK–ENERGY THEOREM

Consider a force, F, applied to a block of mass, m, along the x direction, displacing it a distance d. Assuming that the surface is frictionless (Figure 6.3), the force, F, would then be the only force acting on the block along the x axis. From Newton's second law,

$$F = ma. \tag{6.2}$$

But from the equation of motion

$$v^2 = v_o^2 + 2ax,$$

where the acceleration is

$$a = \frac{v^2 - v_o^2}{2x}.$$

Substituting for a in Equation 6.2 gives

$$F = m\frac{v^2 - v_o^2}{2x},$$

which can be rearranged to give

$$Fx = m\frac{v^2 - v_o^2}{2}.$$

That is,

$$Fx = \frac{1}{2}m(v^2 - v_o^2). \tag{6.3}$$

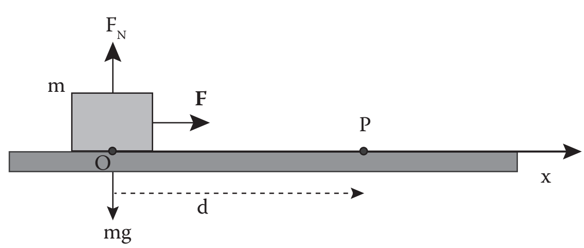


FIGURE 6.3 A block of mass m on a frictionless horizontal surface acted on by a force **F** directed along the displacement **d**. Both are along the x axis.

Defining the kinetic energy, K, for an object of mass, m, moving with a velocity, v, as

$$K = \frac{1}{2}mv^2 \tag{6.4}$$

makes the term on the right hand side of Equation 6.3 recognizable as the difference between the object's final and initial kinetic energies. However, the term on the left hand side is the work, W, done on the block by the force F. This shows that the work done on an object by a force F through a displacement d is equal to the change in its kinetic energy during this displacement. In this argument, neither the weight of the block, mg, nor the force of reaction, F_N, does any work on the block as it moves along the x axis, simply because the angle between the displacement d and each of mg and F_N is 90° and cos 90° = 0. One may rewrite Equation 6.3 in a new form: the work done on an object by the net force acting on it is equal to the change in its kinetic energy, ΔK. That is,

$$W_{net} = \Delta K \tag{6.5a}$$

or

$$W_{net} = \frac{1}{2}m(v^2 - v_o^2). \tag{6.5b}$$

This relation is called *the work–energy theorem*. It establishes the equivalence between work and energy. The unit of energy is Joule.

For a further understanding of the above two sections, Section 6.2 in particular, consider visiting Example 2.8, where a treatment based on energy considerations can be used.

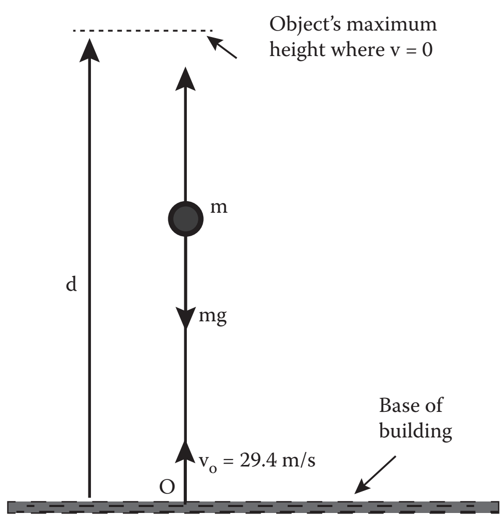
EXAMPLE 6.2

A ball of mass m = 0.100 kg is projected upward from the base of a building with a velocity v_o = 29.4 m/s. Neglecting air resistance, what height does the ball reach before it starts to come down?

Solution

See the figure below. The final velocity of the ball at its highest position is zero. Using the work–energy theorem,

$$W_{net} = \frac{1}{2}m(v^2 - v_o^2).$$



The work done on the ball by its weight during its motion upward is

$$W_{mg} = (mg)d\cos\varphi.$$

Notice that since the gravitational force on the object, mg, is directed downward, while the displacement is upward, the angle then is 180°.

Thus,

$$W_{mg} = (0.100\text{ kg})(9.80\text{ m/s}^2)(d)(\cos 180°) = -(0.980)(d)\text{ J}. \tag{6.6}$$

The change in the ball's kinetic energy is

$$\Delta K = \frac{1}{2}m(v^2 - v_o^2).$$

That is,

$$\Delta K = \frac{1}{2}(0.100\text{ kg})((0.00\text{ m/s})^2 - (29.4\text{ m/s})^2) = -43.2\text{ J}. \tag{6.7}$$

Equating the right sides of (6.6) and (6.7) yields

$$-0.980\,d = -43.2.$$

Solving for d gives

$$d = 44.1\text{ m}.$$

Analysis

1. Checking the above answer with the answer obtained in Example 2.8 part (b) shows that both are the same.
2. Notice that no equations of motion were used in solving the above example. However, care must be taken in defining the following quantities: force(s) acting on the moving object, its displacement, and the angle between them.

EXAMPLE 6.3

Consider the ball in Example 6.2 of mass $m = 0.100$ kg as it falls from its maximum height of 44.1 m. Neglecting air resistance, determine the vertical velocity of the ball just before it hits the ground.

Solution

A sketch of the ball with an initial velocity $v_o = 0.00$ m/s, and with the force of gravity mg acting on it, is shown in the following figure.

Using the work–energy theorem,

$$W_{net} = \frac{1}{2}m(v^2 - v_o^2),$$

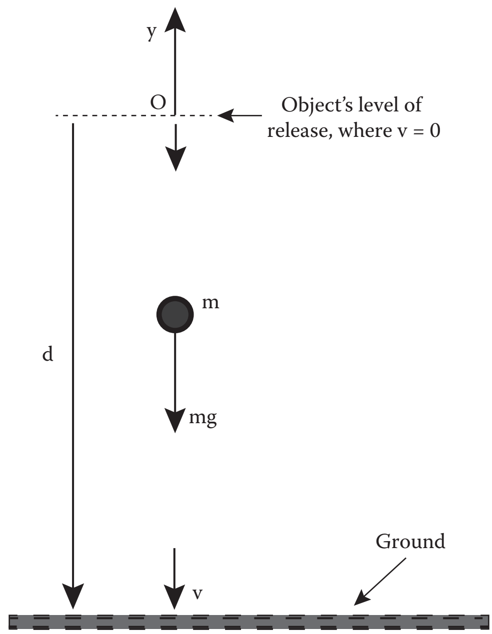
it becomes necessary to find the work done on the ball by the Earth's gravitational force. Notice that since the force is acting downward and the displacement is directed downward, the angle is 0°.

The work done on the ball by its weight during its motion downward is

$$W = (mg)d\cos\varphi.$$

Thus,

$$W = (0.100\text{ kg})\,(9.80\text{ m/s}^2)\,(44.1\text{ m/s})\,(\cos 0°) = 43.2\text{ J}.$$



But from the work–energy theorem

$$W_{net} = \frac{1}{2}m(v^2 - 0.0^2).$$

Equating the two expressions for work gives

$$W_{net} = \frac{1}{2}(0.100 \text{ kg})(v^2 - 0.00^2) = 43.2 \text{ J}.$$

Solving for v gives

$$\mathbf{v} = 29.4 \text{ m/s}.$$

Question

Work out the previous two examples for another ball of mass, m = 0.200 kg. Comment on your answers for the height that you get in Example 6.2 and on the work done on the ball by gravity in the two cases. Check whether there is any effect of mass differences on your result.

Analysis

Notice that the work done on the ball by gravity on its way up (Example 6.2) is the negative of the work done on it on its way down (Example 6.3). Thus, one can conclude that the total work done by gravity on the ball in its round trip is zero.

6.3 CONSERVATIVE FORCES

As described in the analyses following Examples 6.2 and 6.3, the total work done on the object in its round trip is zero. That is,

$$W_{\text{round trip}} = 0.0 \text{ J}.$$

This means that the net change in the object's kinetic energy is zero. In other words, the object, moving upward under gravity only, comes back to the ground's surface with a velocity equal in value to that with which it was launched. Forces that have this property are called conservative forces. Spring forces that will be discussed later in the chapter are also conservative forces. In contrast, forces that do not have this property are labeled as nonconservative forces. The work done on an object by a nonconservative force is always negative, and the net work done on the object in a round trip is also negative. Dissipative forces are good examples of nonconservative forces.

6.4 POTENTIAL ENERGY

From the work–energy theorem, the work done on an object by a force acting on it is equal to the change in its kinetic energy. Thus, from the properties of conservative forces, the kinetic energy of an object is kept undiminished as it returns to that same position if it is under the action of conservative force only. This kinetic energy, a form of ability at a certain position, being restored by the object completely may be explained by assigning to the object a new quantity, called potential energy, that in essence is an alternative notion for "ability." In other words, as the kinetic energy of an object, thrown vertically upwards, decreases as it rises, another quantity, potential energy, designated as U, increases by an amount equal to that lost by the kinetic energy, and when the object reaches its highest position ($v = 0$, $K = 0$), its potential energy assumes a maximum value. This potential energy is what enables the object to come back on its own. While the object is falling, it gains kinetic energy, and when it is back to where it was launched, the kinetic energy is completely regained, whereas the potential energy it previously had gained is completely lost. In symbols, the above discussion can be expressed as

$$\Delta K = -\Delta U. \tag{6.8}$$

The negative sign above is to describe the fact that as kinetic energy decreases, the potential energy increases and vice versa. From Equation 6.8,

$$\Delta K + \Delta U = 0$$

or

$$\Delta(K + U) = 0. \tag{6.9}$$

Since the change in the sum of the K and the U is zero, this sum does not change. That is,

$$(K + U) = \text{constant} \tag{6.10}$$

or

$$(K + U)_i = (K + U)_f = (K + U)_{\text{arbitrary position}}. \tag{6.11}$$

Defining the sum $(K + U)$ as the total mechanical energy, E, that is,

$$E = (K + U). \tag{6.12}$$

Equation 6.11 then becomes

$$E_i = E_f = E_{\text{arbitrary position}}. \tag{6.13}$$

This relation describes the conservation of the total mechanical energy of an object under the influence of a conservative force, or forces, acting on it.

From the work–energy theorem (Equation 6.5a), the work done by any force, or several forces, conservative or nonconservative, is

$$W = \Delta K, \tag{6.14}$$

which, for a conservative force acting on the object, becomes

$$W_C = \Delta K$$

or

$$W_C = -\Delta U. \tag{6.15}$$

This is an important relation stating that the work done on an object by conservative forces is not only associated but equal to a corresponding change in its potential energy with a negative sign imposed on this equality. As stated earlier, the gravitational force, acting on an object, is a good example of a conservative force. As will be discussed in the next section, a spring force, that is, Hooke's force in a spring, is another example of a conservative force.

6.4.1 Potential Energy: Gravity

From Equation 6.15, the potential energy, ΔU, acquired or lost, by an object of mass, m, whose position changes from an initial level to another of a height y is

$$\Delta U = -W_C = -F_{gravity}\ y \cos 180° = -(mg)y\ (-1) = (mg)y. \tag{6.16}$$

This potential energy is positive if y is above the reference level and negative if it is below the reference level. The choice of the reference level is totally optional. Its choice does not affect the solution for the kinematics of the system.

6.4.2 Potential Energy-Elastic Forces

Elastic forces are exerted by rubber bands and springs when they are stretched (rubber bands) and stretched or compressed (springs) by a small amount in comparison with their natural length. These elastic forces are addressed by Hooke's law. It describes an elastic medium as that which when stretched or compressed, within its elastic limit, by an external force, exhibits a self-created force equal in magnitude but opposite in direction to the external force. For a spring, the stretch of the spring exceeding its natural length is proportional to the force applied to it.

Consider a block of mass, m, attached to one end of a spring, while its other end is attached firmly to a solid support (Figure 6.4). An external force needed to elongate the spring or compress it by an amount, x, is

$$F_{EXT} = kx, \tag{6.17}$$

and hence the force developed in the spring is

$$F_{SP} = -kx, \tag{6.18}$$

k is a constant of proportionality, called the spring constant, that describes the stiffness of the spring. It is measured in N/m. The negative sign indicates that the spring force is directed opposite to the displacement.

Therefore, once the external force is removed, the spring force becomes the only force acting on the mass m, in the negative x direction. The spring acts to bring the mass toward its equlbrium

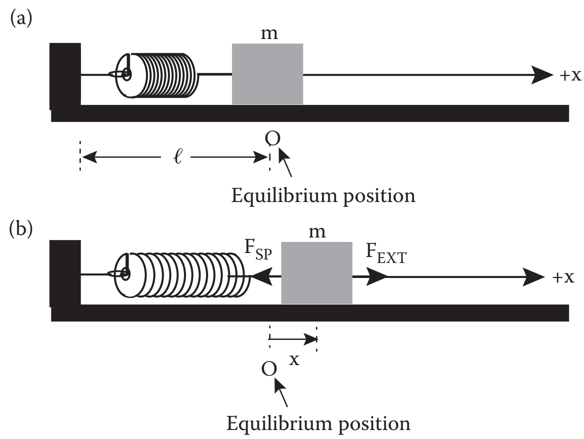


FIGURE 6.4 A block of mass m, attached to one end of a spring whose other end is attached to a solid support, is acted on by an external force F_{EXT} displacing it an amount x.

position. That is why the spring force, F_{SP}, is called a restoring force. The displacement is usually measured from the equilibrium position O, which marks the end of the spring's natural length. The spring's mass here is considered negligible.

To illustrate the concept of the potential energy of a compressed or stretched spring, consider a plot of the magnitude of the force, F_{SP}, versus the displacement, x, as sketched in Figure 6.5. When a spring is stretched to a displacement x from its equilibrium position, O, by an external force, this force has done an amount of work to cause this stretch, and simultaneously a spring force, F_{SP}, is being developed in a direction opposite to the external force. At all instants of stretching the spring, the spring's force is exactly equal to the external force. The work done by the spring's force on mass m is stored as potential energy and is equal to the area under the curve representing the force F versus the position where the elongation has stopped at some value x. The designated area mentioned above is the triangle's area, which is (1/2)(base)(height). Thus,

$$W_C = -\Delta U = -\frac{1}{2}(x)(kx). \tag{6.19}$$

Hence

$$\Delta U = \frac{1}{2}kx^2. \tag{6.20}$$

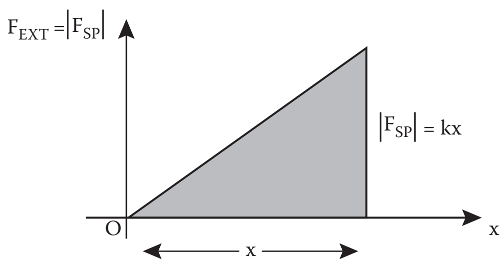


FIGURE 6.5 A plot of the magnitude of the spring force F_{SP} versus the displacement of the block along x, illustrating that the change in the spring potential energy is equal to the area of the shaded triangle of base x and height F_{SP}.

As the total mechanical energy, E, of a system that is acted upon by conservative forces is conserved and is equal to the sum of its kinetic energy, K, and potential energy, U, then the total mechanical energy, E, of the mass-spring system, m, at any arbitrary position, x, from the equilibrium position during the spring motion is

$$E = K + U.$$

That is,

$$E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2. \tag{6.21}$$

6.5 NONCONSERVATIVE FORCES AND THE TOTAL MECHANICAL ENERGY

It has been established from the work–energy theorem that the net work done on an object is equal to the change in its kinetic energy (see Equation 6.5a). That is,

$$W_{net} = \Delta K. \tag{6.22}$$

When several forces, some of which may be conservative and some nonconservative, act on the object, the left hand side of Equation 6.22 can be split into two terms, one, labeled as W_C, for the net work done by the conservative forces and another, W_{NC}, for work done by the nonconservative forces. Accordingly,

$$W_C + W_{NC} = \Delta K.$$

But since from Equation 6.15, $\Delta U = -W_C$, the above relation becomes

$$-\Delta U + W_{NC} = \Delta K$$

or

$$W_{NC} = \Delta K + \Delta U = \Delta(K + U),$$

which, using Equation 6.12 for the sum (K + U), reduces to

$$W_{NC} = \Delta E \tag{6.23}$$

or

$$W_{NC} = E_f - E_i. \tag{6.24}$$

As can be observed from the above relation, the work done by the nonconservative forces is equal to the change in the object's total mechanical energy. Therefore, in contrast to the work W_C done by a conservative force, which is related to the negative change in the object's potential energy, the work W_{NC} done on it by nonconservative forces is equal to the change in its total mechanical energy. Since W_{NC} is always negative, it is clear that E_f for such an object is always less than E_i. This is a direct and sensible result because nonconservative forces such as friction and air resistance are dissipative forces, which, if present, waste part of the total mechanical energy of the object. In other words, for a system under the action of nonconservative forces, the total mechanical energy, E, is not conserved.

6.6 POWER

In the course of defining work, the period of time during which an external force acted on an object was not considered. It is, however, useful in many circumstances to determine the time rate of work done by a force. This brings into effect a new term, called, power, P, delivered to the object by that force. The average power $\bar{P}$ defined as the work done per unit time is expressed mathematically as follows:

$$\bar{P} = \frac{\Delta W}{\Delta t}. \tag{6.25}$$

Since the amount of work, ΔW, done by a constant force F in moving an object through a displacement, Δx, is

$$\Delta W = F\Delta x,$$

Equation 6.25 becomes

$$\bar{P} = \frac{F\,\Delta x}{\Delta t}$$

or

$$\bar{P} = F\bar{v}, \tag{6.26}$$

where $\bar{v}$ is the average velocity of the object during that displacement. The instantaneous power delivered by the force F, however, is defined as

$$P = \lim_{\Delta t \to 0}\left(\frac{\Delta W}{\Delta t}\right),$$

which, from the definition of the instantaneous velocity $v = \lim_{\Delta t \to 0}(\Delta x/\Delta t)$, can be written as

$$P = Fv. \tag{6.27}$$

The unit of power is the Watt (W = J/s).

In ordinary experiences in which this unit is mostly used such as in reporting the power of electrical power plants, generators, motors, and pumps, this unit is rather small, and its multiples such as the kilowatt (kW) and megawatt (MW) are of more use. These quantities in Watts are

$$1\text{ kW} = 10^3\text{ W}$$

and

$$1\text{ MW} = 10^6\text{ W}.$$

From the definition of the Watt as a unit of power, we can see that the unit of energy, J, is equal to W · s. That is,

$$1\text{ J} = 1\text{ W}\cdot\text{s}.$$

Accordingly,

$$1\ \text{kW} \cdot 1\ \text{h} = (1000\ \text{J/s})(3600\ \text{s}).$$

Thus,

$$1\ \text{kW} \cdot \text{h} = 3.6 \times 10^6\ \text{J}.$$

PROBLEMS

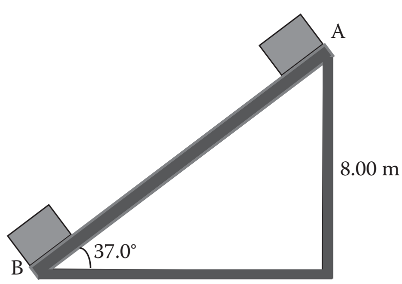
6.1 Two men, Sam and Carl, of 55.0 kg each, are jogging with a constant speed on a track. Sam's speed is 3.00 m/s and Carl's is 6.00 km/s.
   a. Calculate the kinetic energy of each.
   b. Compare between the two kinetic energies by calculating their difference.

6.2 A mother, Mary 62.0 kg, and her daughter Amy, 31.0 kg, are jogging on a track in a morning exercise with the same constant speed of 4.20 m/s.
   a. Calculate the kinetic energy of each.
   b. Compare between the two kinetic energies by calculating the difference between them.

6.3 A 5.00-kg block initially at rest started to move on a frictionless surface under the action of constant force F = 15.0 N, acting parallel to the surface. Considering the first 24.0 m displacement of the block's motion,
   a. Determine the work done on the object by F during its displacement.
   b. Determine the velocity of the block at the 24th meter mark using the work–energy theorem.
   c. Use Newton's second law to find the acceleration of the object during its motion.
   d. Use the value of the velocity found in (c) and pure kinematics to determine the displacement of the object during this motion.

6.4 Jack, 72.0 kg, and his girl friend, Ashley, 58.0 kg, have climbed a mountain of 625 m high.
   a. Calculate the work done on both Jack and Ashley by gravity.
   b. Calculate the gain in their potential energy.
   c. How does the change in potential energy of Jack and Ashley relate to the work done on each by gravity?

6.5 A block of mass m = 2.40 kg was left to slide down on a smooth incline of 8.00 m high and 37.0° angle above the horizontal (the figure below). Calculate
   a. The work done by gravity on the block as it moves down the incline from its initial point at A to its final point at B
   b. The change in the block's potential energy
   c. The change in the block's kinetic energy



6.6 In an accidental push to a small toy box of 1.40 kg, a 6-year-old child sets it in a straight line motion on a slightly rough surface of a tiled floor. The box is slowed down at an average rate of 0.900 m/s^2 till it stops after moving 5.14 m from its initial position.

a. Determine the work done on this box by the force of friction.

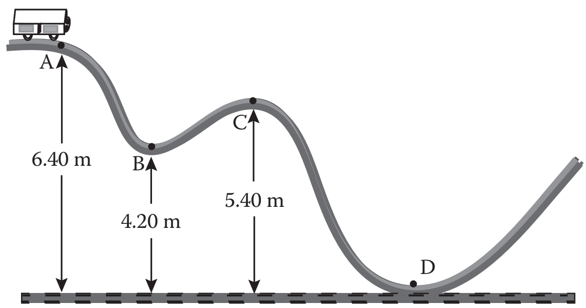
b. Find the force of friction acting on the box during its motion.

6.7 A small car of 6.0×10^2 kg was moving along a straight highway with a velocity of 21.0 m/s when the driver noticed a suspicious object in the middle of his lane 66.0 m away. He slammed on the brakes and was able to bring his car to a full stop in time to avoid running over the object.

a. Determine the change in the car's kinetic energy as it stops.

b. Use the work–energy theorem to calculate the force of friction resulting from pressing the brakes.

6.8 The figure below depicts a frictionless roller coaster in an amusement park where a car is at rest at the edge of the downward inclined track at position A. If the car were released at A, what would be the velocity of the cart at positions B, C, and D?



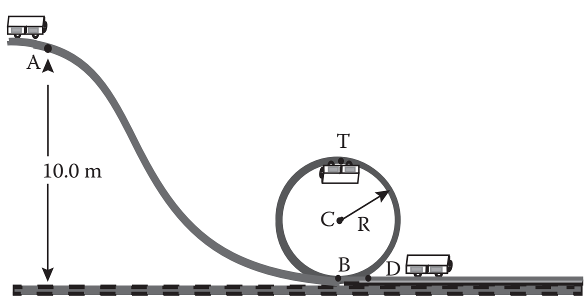
6.9 In a park, a roller coaster car weighing 446 kg, load included, is released from the top start at A of the track shown in the figure below. Assuming the track is frictionless, determine

a. The speed of the car at position B

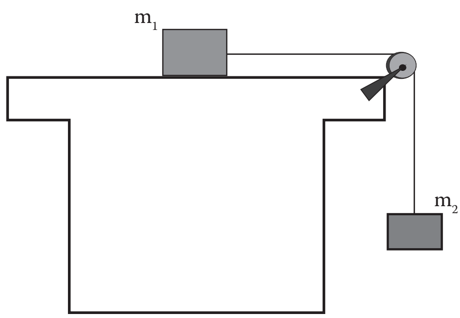
b. The maximum value of the radius R of the loop part that the car can complete without leaving it

c. The speed of the car at its top point T

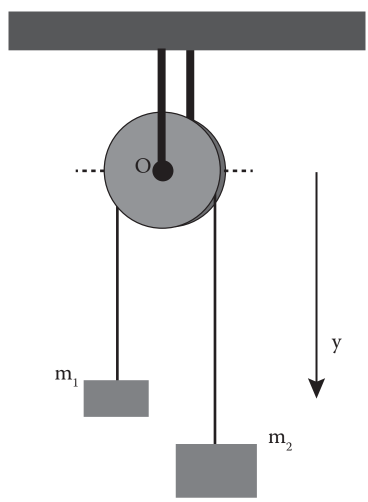
d. The speed of the car as it emerges from the circular track at D that immediately follows position B, but shown separated from it just for clarity



6.10 In the figure below, a block of mass $m_1 = 5.00$ kg resting on a smooth table is connected to another block of mass $m_2 = 4.00$ kg via a string that passes over a smooth pulley at the end of the table. If m_2, initially at rest, started to move downward, calculate the velocity and the acceleration of each block after m_2 moves downward a distance of 2.20 m.

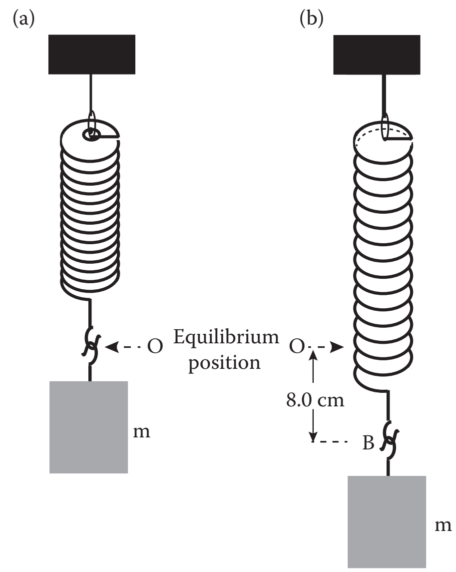


6.11 The figure below shows the Atwood machine that was handled in one of the problems in Chapter 4, where two masses, $m_1 = 4.00$ kg and $m_2 = 8.00$ kg, are connected via a string that passes over a frictionless pulley. The pulley is attached to the ceiling through an axis that passes through its center. Assuming that the string is long enough, use the energy conservation principle to find the velocity of mass m_2 after it has traveled 1.20 m below its initial position.



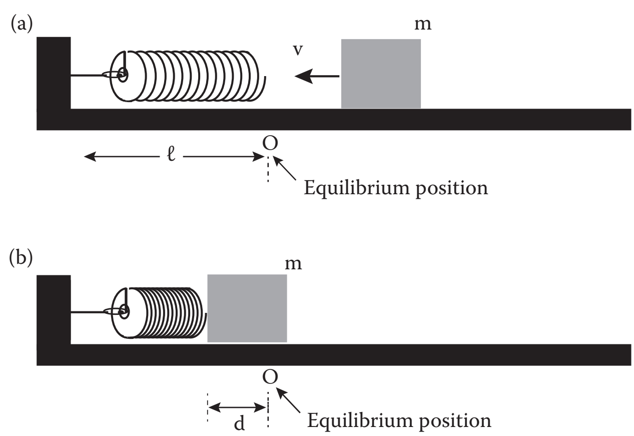
6.12 A block of mass $m = 0.220$ kg, suspended from a spring of spring constant $k = 15.0$ N/m (the following figure), is pulled down a distance 8.00 cm before it was released. Calculate the kinetic, potential, and total energies of the mass-spring system in its first cycle when the block is at

a. The lowest position B
b. A position 2.00 cm below the equilibrium position O



6.13 A block of mass m = 0.450 kg was suddenly pushed to slide on a smooth surface of a counter top colliding with a spring whose other end is attached to a vertical wall (the figure below, part (a)). The block hits the spring with a velocity of 2.40 m/s compressing it a distance d toward the wall (the figure below, part (b)). Knowing that the stiffness constant of the spring is 36.0 N/m, determine

a. The change in the block's kinetic energy as it gets to a full momentary stop
b. The maximum compression distance d of the spring



6.14 A constant force of 7.81 N acts on an object moving with a velocity of 6.40 m/s. Determine the average power that the force is exerting on this object.

6.15 A motor cycle of 28.0 kg accelerated from rest to a speed of 16.0 m/s in 4.20 s. Calculate

a. The average power exerted on this motorcycle during the 4.20-s period
b. The instantaneous power of the motorcycle at the instant t = 2.20 s

7 Linear Momentum and Collision

The definition of linear momentum for a single particle was defined in Chapter 4 (Equation 4.2, Section 4.2.2). In this chapter, the linear momentum for a system of particles will be defined. A new concept known as "center of mass" of such system will be introduced. Newton's laws apply to a system of particles and the motion of its center of mass.

Collisions involve a system of two or more objects. The treatment of a system of particles has special relevance to collisions during which forces of contact act on the colliding objects. It is this force of contact on each object that is called the force of collision. In general, the duration of the forces of collision is rather small compared with the kinematic history of the colliding objects prior to or after the collision, and it is in this short period of time that changing in the objects' velocities in magnitude, direction, or both occurs. In a collision of two objects, the force of one object on the other can be envisioned as a force of action, and hence a force of reaction on it from the other object is simultaneously created. The force of collision on either object is large compared with other external forces, so external forces that may be acting on the object during collision are neglected.

7.1 SYSTEMS OF PARTICLES

In this section, special attention will be given to the center of mass of a system of particles that consists of discrete objects at specific fixed positions. A collection of particles, m_1, m_2, m_3,…, that are positioned in the xy plane (Figure 7.1) has center of mass of coordinates x_{cm} and y_{cm} given by

$$x_{cm} = \frac{m_1x_1 + m_2x_2 + m_3x_3 + \cdots}{(m_1 + m_2 + m_3 + \cdots)}, \tag{7.1a}$$

$$y_{cm} = \frac{m_1y_1 + m_2y_2 + m_3y_3 + \cdots}{(m_1 + m_2 + m_3 + \cdots)}. \tag{7.1b}$$

In the above equations, x_1, x_2, x_3,… and y_1, y_2, y_3,… are the x and y components of the positions r_1, r_2, r_3,… of the objects m_1, m_2, m_3,….

Equations 7.1a and 7.1b can be rewritten as

$$x_{cm} = \frac{m_1x_1 + m_2x_2 + m_3x_3 + \cdots}{M} = \frac{\Sigma_i m_i x_i}{M}, \tag{7.2a}$$

$$y_{cm} = \frac{m_1y_1 + m_2y_2 + m_3y_3 + \cdots}{M} = \frac{\Sigma_i m_i y_i}{M}, \tag{7.2b}$$

where

$$M = m_1 + m_2 + m_3 + \cdots. \tag{7.2c}$$

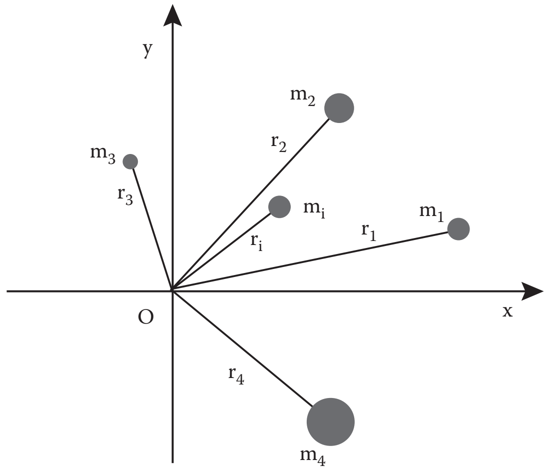
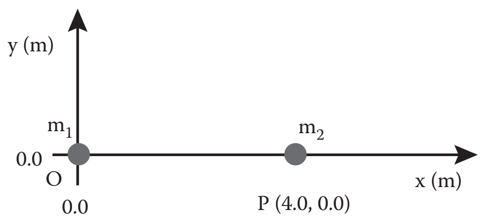


FIGURE 7.1 A system consisting of a collection of discrete particles.

EXAMPLE 7.1

The figure below shows two equal masses, 0.20 kg each, where one is placed at the origin and the other at point P (4.0, 0.0). Find the coordinates of the center of mass of the two masses.



Two equal masses are depicted along the x axis.

Solution

Using Equations 7.1 gives

$$x_{cm} = \frac{m_1x_1 + m_2x_2}{(m_1 + m_2)} = \frac{(0.20\text{ kg})\,(0.00\text{ m}) + (0.20\text{ kg})(4.0\text{ m})}{(0.20\text{ kg} + 0.20\text{ kg})} = 2.0\text{ m}$$

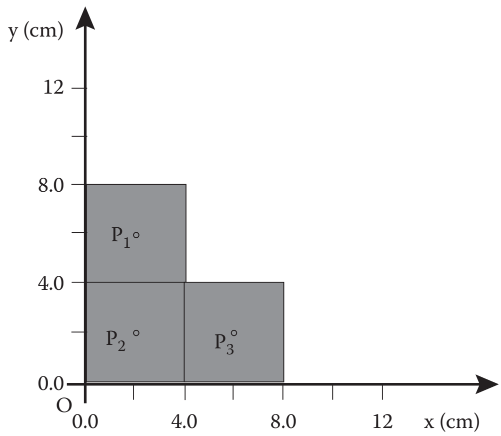
$$y_{cm} = \frac{m_1y_1 + m_2y_2}{(m_1 + m_2)} = \frac{(0.20\text{ kg})\,(0.00\text{ m}) + (0.20\text{ kg})\,(0.00\text{ m})}{(0.20\text{ kg} + 0.20\text{ kg})} = 0.00\text{ m}.$$

Analysis

Since the two masses are equal, the position of the center of mass is expected to be at their midpoint, consistent with what was obtained.

EXAMPLE 7.2

The figure below shows a system of three uniform thin flat squares of a metal, each of mass m = 0.060 kg and of 4.0 cm on a side. The sheets are placed next to each other as shown in the figure. Find the coordinates of the center of mass of this system.



A system of three identical uniform squares.

Solution

As the squares are uniform, each square has a center of mass at the center of the square, and hence these can be treated as three objects each with mass m = 0.060 kg concentrated at the center of the square. The coordinates of the three squares are then at the following positions: P_1 (2.00, 6.00), P_2 (2.00, 2.00), and P_3 (6.00, 2.00).

Using Equations 7.2 gives:

$$x_{cm} = \frac{m_1x_1 + m_2x_2 + m_3x_3 + \cdots}{(m_1 + m_2 + m_3 + \cdots)}$$

$$= \frac{(0.060\text{ kg})\,(2.00\text{ cm}) + (0.060\text{ kg})\,(2.00\text{ cm}) + (0.060\text{ kg})(6.00\text{ cm})}{(0.060\text{ kg} + 0.060\text{kg} + 0.060\text{kg})}$$

$$= \frac{0.60\text{ kg cm}}{(0.18)\text{kg}} = 3.3\text{ cm},$$

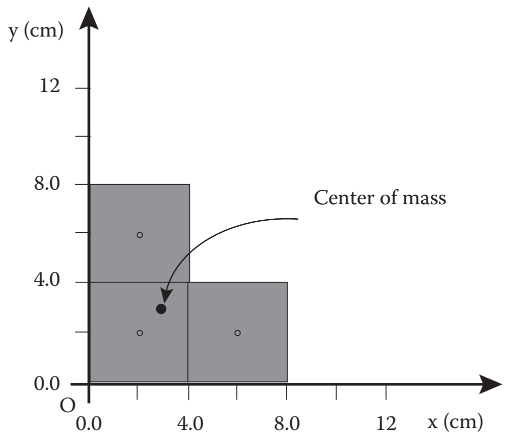
$$y_{cm} = \frac{m_1y_1 + m_2y_2 + m_3y_3 + \cdots}{(m_1 + m_2 + m_3 + \cdots)}$$

$$= \frac{(0.060\text{ kg})\,(6.00\text{ cm}) + (0.060\text{ kg})\,(2.00\text{ cm}) + (0.060\text{ kg})\,(2.00\text{ cm})}{(0.060\text{ kg} + 0.060\text{kg} + 0.060\text{kg})}$$

$$= \frac{0.60\text{ kg cm}}{(0.18)\text{kg}} = 3.3\text{ cm}.$$

Analysis

The coordinates of the center of mass of the system (the following figure) are at P_{cm} (3.30 cm, 3.30 cm), which is a point that lies in the lower left square along the diagonal that makes 45° with the x axis.



The center of mass of a system of three identical uniform squares.

7.2 MOTION OF THE CENTER OF MASS

The significance of the center of mass of a system of particles extends beyond its coordinates to a description about its motion when the system is subjected to external forces. Start with Equations 7.2,

$$x_{cm} = \frac{m_1x_1 + m_2x_2 + m_3x_3 + \cdots}{M},$$

$$y_{cm} = \frac{m_1y_1 + m_2y_2 + m_3y_3 + \cdots}{M}.$$

For a finite displacement in a plane, all particles in the system experience corresponding displacements $\Delta x_1, \Delta x_2, \Delta x_3, \ldots$ and $\Delta y_1, \Delta y_2, \Delta y_3, \ldots$. The system's center of mass experiences displacements

$$\Delta x_{cm} = \frac{m_1\Delta x_1 + m_2\Delta x_2 + m_3\Delta x_3 + \cdots}{M}$$

and

$$\Delta y_{cm} = \frac{m_1\Delta y_1 + m_2\Delta y_2 + m_3\Delta y_3 + \cdots}{M}.$$

The above two equations can be rewritten as

$$M\Delta x_{cm} = m_1\Delta x_1 + m_2\Delta x_2 + m_3\Delta y_3 + \cdots. \tag{7.3a}$$

$$M\Delta y_{cm} = m_1\Delta y_1 + m_2\Delta y_2 + m_3\Delta y_3 + \cdots. \tag{7.3b}$$

The average velocity of the constituent objects and their center of mass can be obtained by dividing Equations 7.3 by Δt. Thus,

$$M\bar{v}_{cm,x} = m_1\bar{v}_{1,x} + m_2\bar{v}_{2,x} + m_3\bar{v}_{3,x} + \cdots, \tag{7.4a}$$

$$M\bar{v}_{cm,y} = m_1\bar{v}_{1,y} + m_2\bar{v}_{2,y} + m_3\bar{v}_{3,y} + \cdots, \tag{7.4b}$$

which as Δt approaches zero gives a corresponding expression for the instantaneous velocity as follows:

$$Mv_{cm,x} = m_1v_{1,x} + m_2v_{2,x} + m_3v_{3,x} + \cdots, \tag{7.5a}$$

$$Mv_{cm,y} = m_1v_{1,y} + m_2v_{2,y} + m_3v_{3,y} + \cdots. \tag{7.5b}$$

A corresponding expression for the average acceleration can be obtained as

$$Ma_{cm,x} = m_1a_{1,x} + m_2a_{2,x} + m_3a_{3,x} + \cdots, \tag{7.6a}$$

$$Ma_{cm,y} = m_1a_{1,y} + m_2a_{2,y} + m_3a_{3,y} + \cdots. \tag{7.6b}$$

From Newton's second law, the above equations give expressions for the force acting on mass M ($= m_1 + m_2 + m_3 + \cdots$) at the center of mass of the system as the sum of the forces acting on all the particles that constitute the system. That is,

$$\mathbf{F}_{cm,x} = \mathbf{F}_{1,x} + \mathbf{F}_{2,x} + \mathbf{F}_{3,x} + \cdots, \tag{7.7a}$$

$$\mathbf{F}_{cm,y} = \mathbf{F}_{1,y} + \mathbf{F}_{2,y} + \mathbf{F}_{3,y} + \cdots. \tag{7.7b}$$

Equations 7.6 and 7.7 describe the motion of the center of mass of the system treating it as if all the mass of the system were concentrated at its center of mass. The equations treat the center of mass as if it were an independent quantity with a mass M.

7.3 LINEAR MOMENTUM OF A SYSTEM OF PARTICLES

Equations 7.5 show that the right hand side is the sum of the linear momenta of all particles in a given direction and the left hand side is the linear momentum of a mass equal to the total mass of the system moving with the velocity of the center of mass of the system in the same direction. That is,

$$\mathbf{P}_{cm,x} = \mathbf{P}_{1,x} + \mathbf{P}_{2,x} + \mathbf{P}_{3,x} + \cdots = \mathbf{P}_x. \tag{7.8a}$$

$$\mathbf{P}_{cm,y} = \mathbf{P}_{1,y} + \mathbf{P}_{2,y} + \mathbf{P}_{3,y} + \cdots = \mathbf{P}_y. \tag{7.8b}$$

Equations 7.7 and 7.8 can be coupled through Newton's second law as follows. Since the rate of change of momentum of a particle (or a system of particles) is equal to the net external force acting on it, the right hand sides of these equations lead to

$$\mathbf{F}_{ext,x} = \frac{\Delta P_x}{\Delta t}, \quad \mathbf{F}_{ext,y} = \frac{\Delta P_y}{\Delta t}.$$

If the net external force on the system is zero, the total momentum of the system is constant. Thus,

$$\mathbf{P}_x = \text{constant}, \quad \mathbf{P}_y = \text{constant}. \tag{7.9}$$

The above equations are the basis for the conservation of linear momentum observed in two-dimensional collisions that will be discussed in the next section.

7.4 COLLISIONS AND CHANGE IN LINEAR MOMENTUM

Consider a collision between two objects. During the collision, each acts on the other by a force, $\mathbf{F}_c$, that will be the only force considered acting on each of the objects. An example of such an event is the collision between a tennis ball and a racket, or between a baseball and a bat, or simply a collision between two balls, two blocks, or two figure skaters accidentally bumping into each other.

Consider a ball of mass m that has been hit by a bat. The force of collision rises from zero at the initial instant of contact to a maximum value, which then declines to zero at the instant of detachment at $t = t_c$ (Figure 7.2a). The exact dependence of the force of collision on time is not trivial to determine. Therefore, it is practical to consider an average value of the collision force $\bar{\mathbf{F}}_c$ that acts on the ball from the very initial instant to the very final instant of collision such that the area under the average force $\bar{\mathbf{F}}_c$ versus time graph would be equal to the area under the exact force acting from 0 to $t = t_c$ (Figure 7.2b).

From Newton's law, the force acting on the ball is

$$\bar{\mathbf{F}}_c = m\mathbf{a} = \frac{m\,\Delta v}{\Delta t} = \frac{m\,(\mathbf{v}_f - \mathbf{v}_i)}{\Delta t}, \qquad (7.10)$$

where $\mathbf{v}_f$ refers to the velocity of the object after collision and $\mathbf{v}_i$ is its velocity prior to collision. The product of mass, m, times the velocity $\mathbf{v}$ defines a new quantity called the linear momentum of the object. The change in the momentum $\Delta\mathbf{P}$ of such an object is

$$\Delta\mathbf{P} = P_f - P_i = m(\mathbf{v}_f - \mathbf{v}_i). \qquad (7.11)$$

Note that $\mathbf{P}$ and $\mathbf{v}$ are vectors. Thus, in collision problems, the proper sign must be affixed to them. Back to Equation 7.10 that, from Newton's second law, becomes

$$\bar{\mathbf{F}}_c = \frac{\Delta\mathbf{P}}{\Delta t}. \qquad (7.12)$$

That is,

$$\Delta\mathbf{P} = \bar{\mathbf{F}}_c\Delta t. \qquad (7.13)$$

A schematic of a collision between two balls of masses m_1 and m_2 in a head on collision is demonstrated in Figure 7.3.

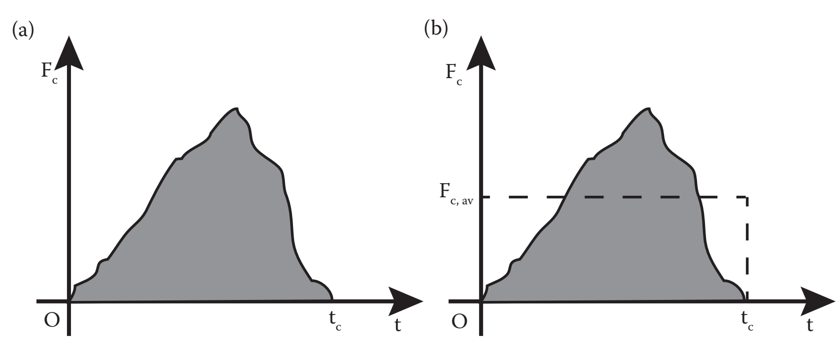


FIGURE 7.2 (a) The force of collision (solid curve) and (b) a rectangle of area equal to that under the solid curve. The rectangle's width represents the average value of the force of collision $\bar{F}_c$ (horizontal dashed line), acting over the same period, $0 \to t_c$.

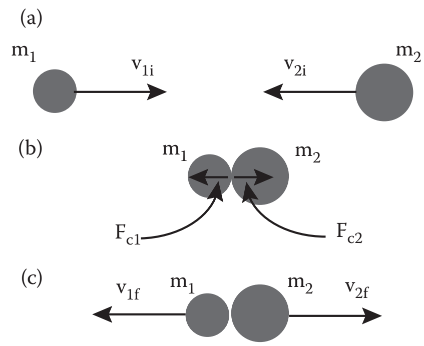


FIGURE 7.3 A head on collision between two balls of different masses: (a) prior to collision, (b) during collision, and (c) right after collision.

Since external forces in a collision are ignored, the system as a whole is under no external force. Therefore, the force with which m_1 acts on m_2 during collision, call it F_{c2}, is simultaneously associated with an equal but opposite force of reaction F_{c1} that m_2 exerts on m_1. That is,

$$\bar{\mathbf{F}}_{c1} = \frac{\Delta \mathbf{P}_1}{\Delta t}, \quad \bar{\mathbf{F}}_{c2} = \frac{\Delta \mathbf{P}_2}{\Delta t}.$$

As

$$\bar{\mathbf{F}}_{c1} + \bar{\mathbf{F}}_{c2} = 0,$$

then

$$\frac{\Delta \mathbf{P}_1}{\Delta t} + \frac{\Delta \mathbf{P}_2}{\Delta t} = 0.$$

Accordingly,

$$\Delta \mathbf{P}_1 + \Delta \mathbf{P}_2 = 0. \tag{7.14}$$

That is,

$$\Delta \mathbf{P} = \mathrm{P}_f - \mathrm{P}_i = 0. \tag{7.15}$$

Here $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2$. This conclusion applies to all colliding systems regardless of the number of objects involved.

Thus,

$$\mathbf{P} = \text{constant}. \tag{7.16}$$

Denoting the momenta just prior to and right after collision by $\mathbf{P}_i$ and $\mathbf{P}_f$, respectively, Equation 7.15 may be written as

$$\mathbf{P}_i = \mathbf{P}_f. \tag{7.17}$$

Equation 7.16 reflects an important principle called conservation of linear momentum, which applies to all colliding objects. For a system consisting of two colliding objects, Equation 7.17 written in terms of their velocities becomes

$$m_1\mathbf{v}_{1i} + m_2\mathbf{v}_{2i} = m_1\mathbf{v}_{1f} + m_2\mathbf{v}_{2f}. \tag{7.18}$$

EXAMPLE 7.3

Two men, Sam and Carl, of 55.0 kg each, are jogging on a track at a constant speed. Sam's speed is 3.00 m/s and Carl's speed is 6.00 m/s.

a. Calculate the momentum of each.
b. Find the ratio of the momentum of Sam to Carl.

Solution

a. The momentum of Sam is

$$P = mv = (55.0 \text{ kg})(3.00 \text{ m/s}) = 1.65 \times 10^2 \text{ kg m/s}.$$

The momentum of Carl is

$$P = mv = (55.0 \text{ kg})(6.00 \text{ m/s}) = 3.30 \times 10^2 \text{ kg m/s}.$$

b. The ratio of the momentum of Sam to Carl is $(1.65 \times 10^2 \text{ kg m/s}/3.30 \times 10^2 \text{ kg m/s}) = 0.500$.

Analysis

Notice that with the masses of both Sam and Carl being the same, the momentum of Carl is twice that of Sam due to Carl's speed being twice that of Sam.

EXAMPLE 7.4

A golf player hits a stationary golf ball (0.450 kg) imparting to it a horizontal speed of 22.0 m/s. If the time of contact between the club and the ball is 2.40 s, calculate

a. The final momentum of the ball immediately after being hit
b. The change in the momentum of the ball due to collision
c. The average force of collision acting on the ball during contact

Solution

a. Labeling the ball as object 1 and the golf club as object 2, the momentum of the ball after being hit is

$$P_{1f} = m\mathbf{v}_{1f} = (0.450 \text{ kg})\ (22.0 \text{ m/s}) = 9.90 \text{ kg m/s}.$$

b. Change in the momentum of the ball due to collision is

$$\Delta\mathbf{P}_1 = P_{1f} - P_{1i} = m(\mathbf{v}_{1f} - \mathbf{v}_{1i}) = (0.450 \text{ kg})(22.0 \text{ m/s} - 0.00 \text{ m/s}) = 9.90 \text{ kg m/s}.$$

c. The average force of collision acting on the ball during contact is

$$\bar{\mathbf{F}}_{c1} = \frac{\Delta\mathbf{P}_1}{\Delta t} = \frac{(9.90 \text{ kg m/s} - 0.00 \text{ kg m/s})}{2.40\ \text{ s}} = 4.13 \text{ N}.$$

Analysis

1. From the discussion leading to Equation 7.14, it can be deduced that the force the club experiences during collision would also be 4.13 N, opposite in direction to that experienced by the ball.
2. Also observe from Equation 7.11 that the change in momentum of the ball due to collision should be equal to the change in the momentum of the club. This follows from the conservation of momentum of the colliding system just prior and just after the collision takes place. This is discussed in examples of the next section.

7.5 IMPULSE

The change in momentum, $\Delta\mathbf{p}$, of an object, that is, the product $\bar{\mathbf{F}}_c\Delta t$, defines another new quantity known as the impulse, **I**. For any object,

$$\mathbf{I} = \Delta P = \bar{\mathbf{F}}_c\Delta t. \tag{7.19}$$

In the above discussion, the impulse imparted to mass m_1 is

$$\mathbf{I}_1 = \bar{\mathbf{F}}_{c1}\Delta t \tag{7.20a}$$

and the impulse imparted to mass m_2 is

$$\mathbf{I}_2 = \bar{\mathbf{F}}_{c2}\Delta t. \tag{7.20b}$$

As can be seen from Equation 7.14,

$$\mathbf{I}_1 + \mathbf{I}_2 = 0,$$

implying that

$$\mathbf{I}_1 = -\mathbf{I}_2. \tag{7.21}$$

As the impulse of an object is equal to the change in its momentum which is a vector, the impulse is then a vector. From the above, it can be concluded that the total impulse of the system is zero. That is,

$$\mathbf{I} = 0. \tag{7.22}$$

Exercise

Determine the impulse of both the ball and the club in Example 7.4.

EXAMPLE 7.5

A bullet of 22.5 g fired from a hand gun experiences, during the firing process, an average force of 2225 N. If the force acting on the bullet during its firing lasts for 2.20×10^{-3} s, what

a. Would be the value of the impulse imparted to the bullet right after firing?
b. Would be the value of the impulse imparted to the gun right after firing the bullet?
c. Would be the bullet's speed right after firing?

Solution

a. Labeling the bullet as object 1 and the gun as object 2, the impulse imparted to the bullet during the firing process is

$$\mathbf{I}_1 = \bar{\mathbf{F}}_{c1}\Delta t = (2225\ \text{N})(2.20 \times 10^{-3}\ \text{s}) = 4.90\ \text{Ns}.$$

b. From Equation 7.21, the impulse imparted to the gun during the firing process is

$$\mathbf{I}_2 = -\,\mathbf{I}_1 = -4.90\ \text{Ns}.$$

The minus sign is just to demonstrate that the impulse of the gun is opposite in direction to that of the bullet. Since the impulse of the bullet is obviously forward, the impulse of the gun should be backward.

c. From Equation 7.19, the impulse I_1 imparted to the bullet during its firing is

$$\mathbf{I}_1 = \Delta\mathbf{P}_1 = 4.90\ \text{Ns}.$$

Thus,

$$m(\mathbf{v}_{1f} - \mathbf{v}_{1i}) = 4.90\ \text{Ns}.$$

Substituting for m and $\mathbf{v}_1$, the above equation becomes

$$(0.0225\ \text{kg})(\mathbf{v}_{1f} - 0) = 4.90\ \text{Ns}.$$

This gives

$$\mathbf{v}_{1f} = 218\ \text{m/s}.$$

Analysis

In part (c), the bullet's muzzle velocity is 218 m/s. This is equal to 490 mi/h, interestingly high and probably unexpected!

7.6 COLLISIONS IN TWO DIMENSIONS

The conclusion arrived at in Equations 7.17 and 7.16 can be generalized to two-dimensional collisions. That is, the linear momentum for a system of two objects is conserved. Since Equation 7.17 is a vector equation, it can be resolved into its components. A typical sketch of a collision between two objects, two billiard balls for example that results in the scattering of the two balls in a plane (the plane here is the surface of the billiard table) that combines their initial and final velocities, is illustrated in Figure 7.4.

Taking the plane of collision to be the xy plane, Equation 7.17 applied to this case implies that

$$\mathbf{P}_x = \mathbf{P}_{xf}, \quad \mathbf{P}_y = \mathbf{P}_{yf}.$$

That is,

$$m_1\mathbf{v}_{1i\,x} + m_2\mathbf{v}_{2i\,x} = m_1\mathbf{v}_{1f\,x} + m_2\mathbf{v}_{2f\,x}, \tag{7.23}$$

$$m_1\mathbf{v}_{1i\,y} + m_2\mathbf{v}_{2i\,y} = m_1\mathbf{v}_{1f\,y} + m_2\mathbf{v}_{2f\,y}. \tag{7.24}$$

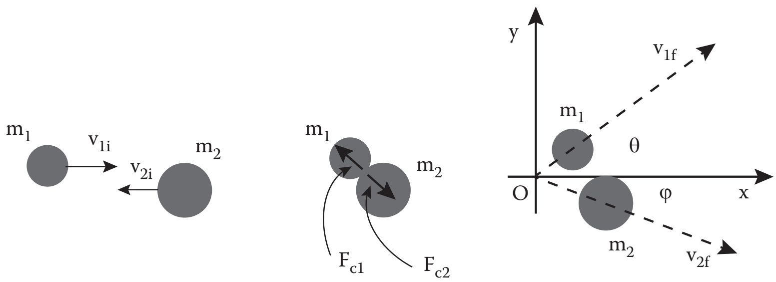


FIGURE 7.4 Collision between two balls on a billiard table. The balls after the collision scatter in the plane of the table.

Of course, if the collision is three dimensional, a third component of the momenta prior to and after collision along the z axis would be needed. That would take the form

$$m_1\mathbf{v}_{1i\,z} + m_2\mathbf{v}_{2i\,z} = m_1\mathbf{v}_{1f\,z} + m_2\mathbf{v}_{2f\,z}. \tag{7.25}$$

In case the collision involves three or more objects, the above equations are then generalized to

$$m_1\mathbf{v}_{1i\,x} + m_2\mathbf{v}_{2i\,x} + m_3\mathbf{v}_{3i\,x} + \cdots = m_1\mathbf{v}_{1f\,x} + m_2\mathbf{v}_{2f\,x} + m_3\mathbf{v}_{3f\,x} + \cdots, \tag{7.26}$$

$$m_1\mathbf{v}_{1i\,y} + m_2\mathbf{v}_{2i\,y} + m_3\mathbf{v}_{3i\,y} + \cdots = m_1\mathbf{v}_{1f\,y} + m_2\mathbf{v}_{2f\,y} + m_3\mathbf{v}_{3f\,y} + \cdots, \tag{7.27}$$

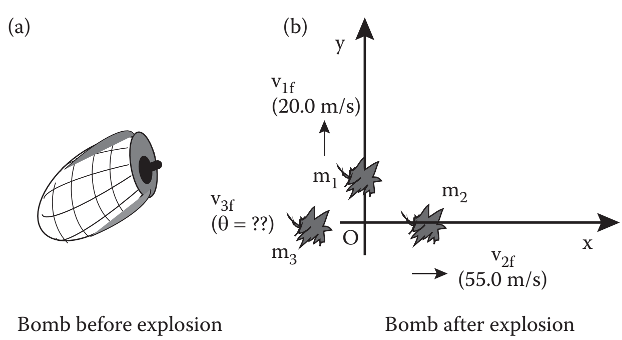
$$m_1\mathbf{v}_{1i\,z} + m_2\mathbf{v}_{2i\,z} + m_3\mathbf{v}_{3i\,z} + \cdots = m_1\mathbf{v}_{1f\,z} + m_2\mathbf{v}_{2f\,z} + m_3\mathbf{v}_{3f\,z} + \cdots. \tag{7.28}$$

EXAMPLE 7.6

A bomb exploded into three equal parts. One of the fragments was seen moving upward with a velocity of 20.0 m/s and another was seen moving to the right with a velocity of 55.0 m/s. Assuming that all fragments had moved in the xy plane, determine the magnitude and direction of the velocity of the third fragment.

Solution

The momentum of this system, the bomb prior to and after explosion, is conserved. Since no fragment was seen to move in the forward or backward direction, that is, none in the z direction, this collision is two dimensional (the figure below), x along the horizontal to the right and y along the vertical direction. Since the three fragments are equal in mass, $m_1 = m_2 = m_3 = m$.



A bomb that explodes into three fragments, all moving in the xy plane.

Applying Equation 7.26,

$$m_1\mathbf{v}_{1i\,x} + m_2\mathbf{v}_{2i\,x} + m_3\mathbf{v}_{3i\,x} + \cdots = m_1\mathbf{v}_{1f\,x} + m_2\mathbf{v}_{2f\,x} + m_3\mathbf{v}_{3f\,x} + \cdots.$$

Thus,

$$0.00 = m\,(55.0) + m(0.00) + m\mathbf{v}_{3f\,x} + \cdots.$$

This gives

$$\mathbf{v}_{3f\,x} = -(55.0)\text{ m/s}.$$

Similarly, for the y direction, Equation 7.27

$$m_1\mathbf{v}_{1i\,y} + m_2\mathbf{v}_{2i\,y} + m_3\mathbf{v}_{3i\,y} + \cdots = m_1\mathbf{v}_{1f\,y} + m_2\mathbf{v}_{2f\,y} + m_3\mathbf{v}_{3f\,y} + \cdots$$

becomes

$$0.00 = m(0.00) + m(20.0) + m\mathbf{v}_{3f\,y}.$$

This gives

$$\mathbf{v}_{3f\,y} = -20.0\text{ m/s}.$$

The third fragment is in the xy plane with two components, one along the negative x direction (i.e., to the left) and another in the negative y direction (i.e., downward). The magnitude of the velocity of this third fragment is

$$\mathbf{v}_{3f} = \sqrt{(v_{3f\,x}^2) + (v_{3f\,y}^2)} = \sqrt{(-55.0)^2 + (-20.0)^2} = 58.5\text{ m/s}.$$

The direction of the velocity of the third fragment is obtained from

$$\theta = \tan^{-1}\left(\frac{\mathbf{v}_{3f\,y}}{\mathbf{v}_{3f\,x}}\right) = \tan^{-1}\left(\frac{-55.0\text{ m/s}}{-20.0\text{ m/s}}\right) = \tan^{-1} 2.75 = 70°.$$

As the two components are in the third quadrant, this angle corresponds to an angle of (180° + 70°) = 250° with respect to the positive x axis.

7.7 TYPES OF COLLISION

The kinetic energy of a system of colliding objects before and after a collision is employed as a criterion to categorize the type of collision they encounter. In this regard, there are two types of collision:

7.7.1 Elastic Collision

In addition to the conservation of linear momentum that applies to all collisions, the conservation of kinetic energy of the colliding objects just before and just after collision also holds for elastic collisions. Therefore, for an elastic collision

$$K_{i1} + K_{2i} = K_{1f} + K_{2f}.$$

In terms of the masses and velocities of the colliding objects, the above relation becomes

$$\frac{1}{2}m_1\mathbf{v}_{1i}^2 + \frac{1}{2}m_2\mathbf{v}_{2i}^2 = \frac{1}{2}m_1\mathbf{v}_{1f}^2 + \frac{1}{2}m_2\mathbf{v}_{2f}^2. \tag{7.29}$$

It can be shown that for a one-dimensional collision of two objects, the above condition, combined with the conservation of linear momentum, leads to the following result:

$$v_{1i} - v_{2i} = -(v_{1f} - v_{2f}). \tag{7.30}$$

7.7.2 Inelastic Collision

In this kind of collision, the kinetic energy of a system just before and just after collision is not conserved. Thus, for a system of two colliding objects, that is,

$$K_{1i} + K_{2i} \neq K_{1f} + K_{2f},$$

$$(K_{1f} + K_{2f}) - (K_{1i} + K_{2i}) \neq 0, \tag{7.31}$$

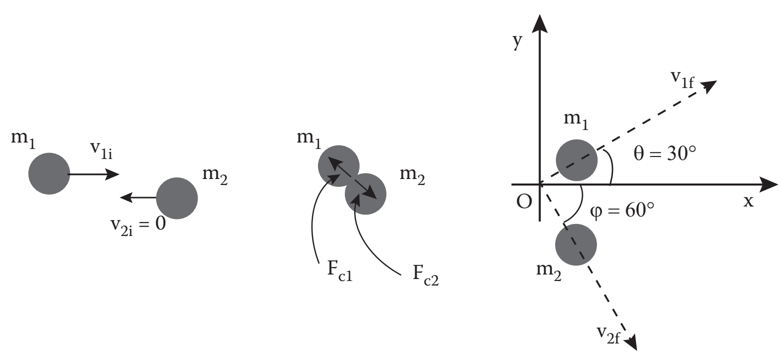
or

$$\Delta\,(K)_{system} \neq 0.$$

This difference in the kinetic energy of the system can appear as heat or deformation of the colliding objects.

EXAMPLE 7.7

A small marble sphere was rolled horizontally across a smooth table top with a velocity of 8.0 m/s and hits, in a slightly an off-center collision, a stationary identical sphere (the figure below). After collision, the first sphere moved along the line that makes an angle of 30° with the forward direction while the second sphere moved along the line that makes an angle of 60° with the forward direction as shown. Assuming that the collision was elastic, determine the magnitude of the final velocity of the second sphere.



Solution

As the two marbles are identical, let $m_1 = m_2 = m$.

From Equation 7.23

$$m_1\mathbf{v}_{1i\,x} + m_2\mathbf{v}_{2i\,x} = m_1\mathbf{v}_{1f\,x} + m_2\mathbf{v}_{2f\,x}.$$

Thus,

$$m(8.0\ \text{m/s}) + m(0.0) = m\mathbf{v}_{1f}\cos 30° + m\mathbf{v}_{2f}\cos 60°$$

or

$$(8.0\ \text{m/s}) - \mathbf{v}_{1f}(0.87) = (0.50)\mathbf{v}_{2f}. \tag{7.32}$$

Similarly from Equation 7.24

$$m_1\mathbf{v}_{1i\,y} + m_2\mathbf{v}_{2i\,y} = m_1\mathbf{v}_{1f\,y} + m_2\mathbf{v}_{2f\,y}.$$

Thus,

$$m(0.0\text{ m/s}) + m(0.0\text{ m/s}) = m\mathbf{v}_{1f}\sin 30° - m\mathbf{v}_{2f}\sin 60°$$

or

$$0.0\text{ m/s} = m\mathbf{v}_{1f}\sin 30° - m\mathbf{v}_{2f}\sin 60°.$$

Thus,

$$(0.50)\mathbf{v}_{1f} = \mathbf{v}_{2f}\,(0.87)$$

or

$$\mathbf{v}_{1f} = \mathbf{v}_{2f}(1.74). \tag{7.33}$$

Substituting Equation 7.33 in Equation 7.32 gives

$$(8.0\text{ m/s}) - (1.74\mathbf{v}_{2f})(0.87) = (0.50)\mathbf{v}_{2f}$$

or

$$\mathbf{v}_{2f} = 4.0\text{ m/s}.$$

Substituting for $\mathbf{v}_{2f} = 4.0$ m/s in Equation 7.33 yields

$$\mathbf{v}_{1f} = \mathbf{v}_{2f}(1.74) = 6.9\text{ m/s}.$$

Analysis

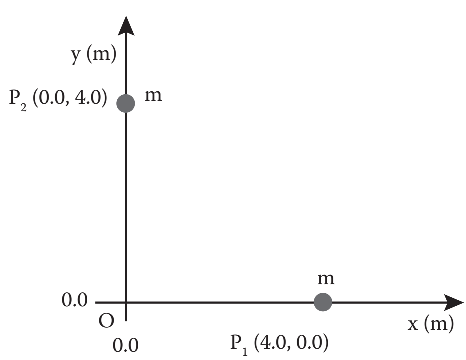
In the above example, the off-center collision influences the amount the two marbles get scattered and their velocities after collision.

PROBLEMS

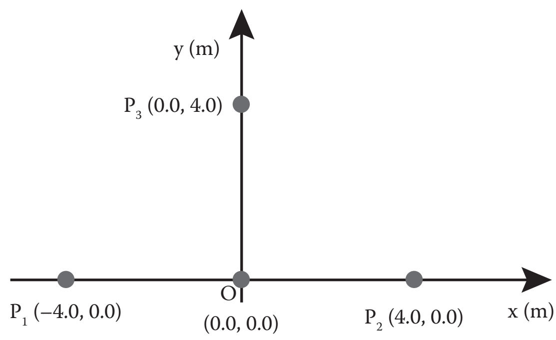
7.1 The figure below shows a system of two equal masses, 0.20 kg each, placed at the following positions: P_1 (4.0, 0.0) and P_2 (0.0, 4.0).

a. Using the mathematical definition of the center of mass, find its x and y coordinates.

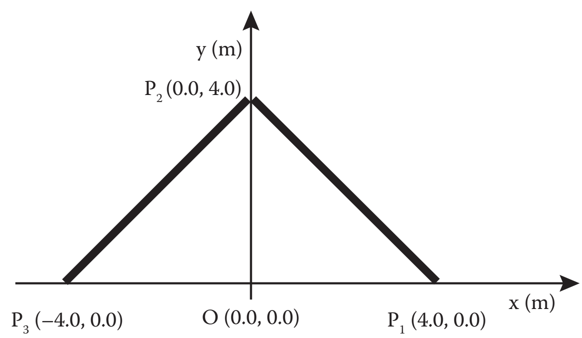
b. Use symmetry to find the coordinates of the center of the above two particles.



7.2 The figure below shows a system of three equal masses, 0.20 kg each, placed at the following positions: P_1 (−4.0, 0.0), P_2 (4.0, 0.0), and P_3 (0.0, 4.0). Find the coordinates of the center of mass of this system.



7.3 The figure below shows a system of two thin uniform rods of equal masses, 0.25 kg each. The rods, welded seamlessly at one end, form a rigid system oriented as shown in the figure. Find the coordinates of its center of mass.

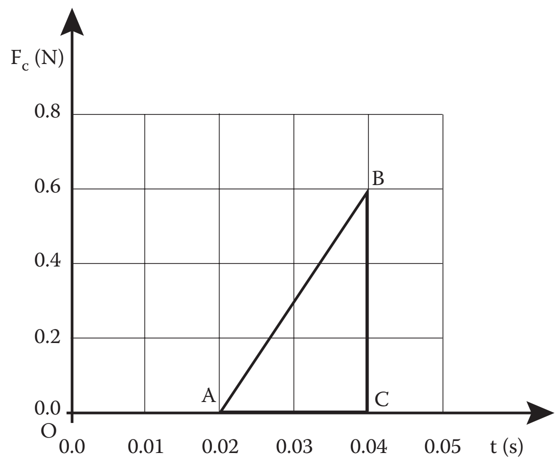


7.4 Mary, 60.0 kg, and her daughter Susan, 40.0 kg, are jogging on a track with the same constant speed of 5.0 m/s.
   a. Calculate the momentum of each.
   b. Find the ratio of the momentum of Mary to Susan.

7.5 A car of mass m = 1100 kg, moving with a velocity of 22.0 m/s (49 mi/h), skidded off the road hitting a tree that brought the car to rest in 2.0 s. Determine
   a. The momentum of the car prior to collision
   b. The average force that acted on the car during collision
   c. The average force that acted on the tree during collision
   d. The impulse imparted to the tree

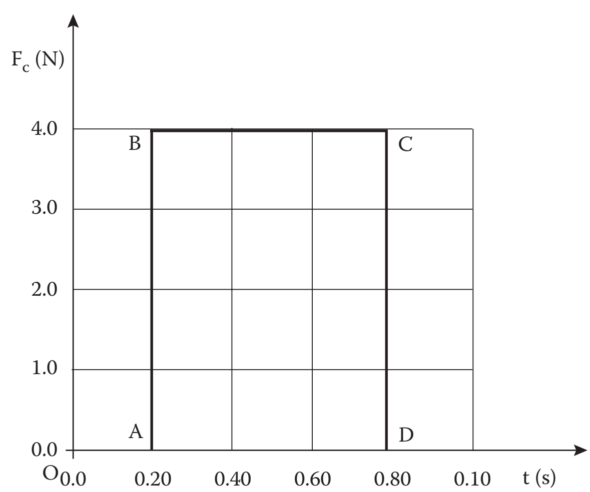
7.6 The following figure shows a horizontal force of collision as a function of time acting on a 0.400-kg object for a duration of 0.02 s between its initial and final instants labeled as A and C, respectively.
   a. Find the magnitude of the impulse of the object.
   b. Assuming that the object is initially at rest, find the magnitude and direction of its final velocity.

c. Find what constant force would act on the object between the very initial and final instants of collision and result in the same impulse produced by F_c. Sketch this force as a function of time.



7.7 A 0.200-kg ball moving with 4.00 m/s in the negative x direction is hit by a bat that exerted on it an average force of 4.0 N in the positive x direction (The figure below). The time of contact between B and C is 0.60 s.

a. Find the magnitude of the impulse of the ball.
b. Find the magnitude and direction of the ball's final velocity.



7.8 A rubber ball of 215 g mass fell from a boy's hand hitting the floor with a velocity v = 2.45 m/s. The ball bounced vertically back with a velocity of the same value after a contact of 0.42 s with the floor.

a. Calculate the impulse of the ball.
b. Calculate the average force with which the floor acted on the ball during the contact.

7.9 A bullet of mass m = 14.0 g, fired horizontally with a speed v = 280 m/s at a tree, came to a halt after penetrating 16.2 cm into the tree. Calculate

a. How long it took the bullet to come to a halt?
b. The average force of collision acting on the bullet

7.10 In a hunting season, a man of mass 85.0 kg, holding a 3.6-kg rifle against his shoulder, fires a 6.00-g bullet at his target. If the muzzle velocity of the bullet is 168 m/s, and treating the rifle and the bullet as an isolated system,

a. Determine the recoil velocity of the rifle.
b. Find the recoil velocity of the man and rifle combined.

7.11 A bullet of mass m = 14.0 g, fired horizontally with a speed v = 288 m/s, hit a wooden block of mass m = 5.08 kg placed on a flat large rough table counter. The block and bullet embedded in came to a halt after moving across the counter a distance of 3.0 m.

a. Calculate the speed of the block immediately after collision.
b. The force of friction between the block and the counter's top.

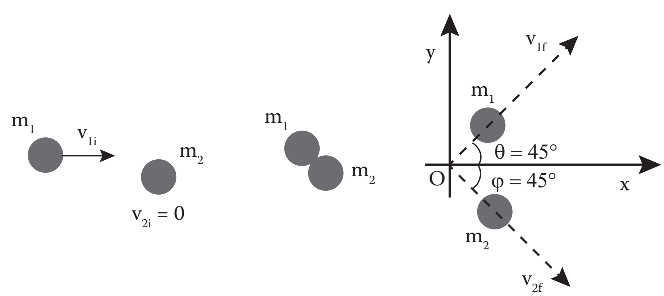
7.12 A small marble sphere, rolled across a smooth table top with a velocity of 8.0 m/s, has a head on elastic collision with a stationary identical sphere. What would be the velocities of the two spheres immediately after collision?

7.13 In a head on elastic collision, a marble of mass M at rest is struck by another marble of mass m moving with a velocity v = 2.0 m/s. Assuming that M = 2 m, determine the speed of each marble immediately after the collision.

7.14 A truck of mass m_1 = 3600 kg with a velocity v_1 = 32 m/s traveling east and a small car of mass m_2 = 550 kg with a velocity v_2 = 22 m/s moving north collided at an intersection. If the car became attached to the truck and both moved together as one wreck along at an angle θ northeast, determine

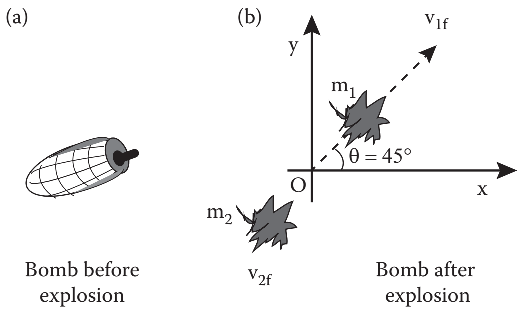
a. The velocity of the two immediately after collision
b. The direction along which the wreck will move
c. The change in the kinetic energy of the two vehicles due to contact
d. Whether or not the collision was elastic

7.15 A small marble sphere was rolled horizontally across a smooth table top with a velocity of 12.0 m/s and hits, in a slightly an off-center collision, a stationary identical sphere (the figure below). After collision, the first sphere was observed moving in a direction of 45° with the forward northeast direction and the second moved in a direction that makes an angle of 45° with the forward southeast direction. Determine the magnitude of the final velocity of the two spheres.



7.16 A bomb exploded into two equal parts. One of the fragments was seen moving north with a velocity of 18.0 m/s. Determine the magnitude and direction of the velocity of the second fragment.

7.17 Repeat the previous problem assuming that one of the fragments was seen moving northeast with a velocity of 12.7 m/s (the figure below). Determine the magnitude and direction of the velocity of the second fragment.



8 Rotational Motion

Translational motion and circular motion, uniform and nonuniform of an object, were discussed in Chapters 3 and 4. In this chapter, rotational motion of a point or extended object will be introduced. A force acting on an extended object creates a torque that rotates it about a fixed axis. A solid object of finite physical size is known as a rigid body. The axis of rotation could be about the center of mass of the rigid body or about other points where it is pivoted. In this chapter, Sections 8.2 through 8.7 discuss the kinematics of a rigid body, and Section 8.8 reviews the dynamics of the rigid body. In addition to drawing a parallel between Newton's laws applied to point-like objects in linear motions and those applied to rotational motion of a rigid body, translational and rotational motions of a rigid body are discussed in detail.

8.1 ANGULAR KINEMATIC QUANTITIES

Consider a point-like object moving in a circle of radius r (Figure 8.1). As the object moves from point P_1 at t_1 to point P_2 at t_2, it sweeps through an arc Δs that subtends an angle $\Delta\theta$ ($= \theta_2 - \theta_1$) at the center, in a time interval $\Delta t = t_2 - t_1$. From geometry

$$\Delta s = r\Delta\theta, \tag{8.1}$$

where $\Delta\theta$ is an angular displacement, measured in radians. To convert an angle expressed in degrees to radians, the following relation may be used:

$$\theta_{rad} = \left(\frac{2\pi rad}{360_{deg}}\right)\theta_{deg}, \tag{8.2a}$$

and from radians to degrees, the conversion is

$$\theta_{deg} = \left(\frac{360}{2\pi}\right)\theta_{rad}. \tag{8.2b}$$

Dividing Equation 8.1 by Δt gives the average linear velocity $\bar{v}$; that is,

$$\frac{\Delta s}{\Delta t} = r\frac{\Delta\theta}{\Delta t}$$

or

$$\bar{v} = r\bar{\omega}. \tag{8.3}$$

For an object experiencing a constant linear acceleration,

$$\bar{v} = \frac{v_1 + v_2}{2}.$$

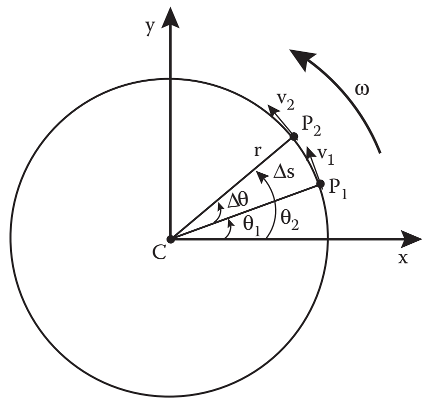


FIGURE 8.1 A point like object moving in a circle of radius r.

Here, v_1 and v_2 are the instantaneous tangential, or linear, velocities of the object at points P_1 and P_2, respectively. Similarly, $\bar{\omega}$ is the average angular velocity given by

$$\bar{\omega} = \frac{\omega_1 + \omega_2}{2}, \tag{8.4}$$

where ω_1 and ω_2 are the instantaneous angular velocities of the object at points P_1 and P_2, respectively. Equation 8.4 holds only if the change in the object's angular velocity is uniform.

8.2 ROTATIONAL MOTION IN A PLANE

Equation 8.3 can be used to establish a connection between the instantaneous linear velocity v and the angular velocity ω at any point on the circle as

$$v = r\omega. \tag{8.5}$$

The average tangential, or linear, acceleration of the object during its motion between points P_1 and P_2 is defined as

$$\bar{a} = \frac{v_2 - v_1}{t_2 - t_1}. \tag{8.6}$$

Using Equation 8.5 in Equation 8.6 gives

$$\bar{a} = \frac{r\,(\omega_2 - \omega_1)}{(t_2 - t_1)}$$

or

$$\bar{a} = r\bar{\alpha}, \tag{8.7}$$

where

$$\bar{\alpha} = \frac{(\omega_2 - \omega_1)}{(t_2 - t_1)} \tag{8.8a}$$

is the average angular acceleration. If the angular acceleration is constant, $\bar{\alpha} = \alpha$, then the above equation can be expressed as

$$\alpha = \frac{(\omega_2 - \omega_1)}{(t_2 - t_1)}.$$

Setting ω for ω_2 as the object's final angular velocity, ω_1 for ω_o as its initial angular velocity, t for t_2, and setting $t_1 = 0$, the above equation becomes

$$\alpha = \frac{\omega - \omega_o}{t}. \tag{8.8b}$$

For a constant angular acceleration, the tangential acceleration is also constant and is given by

$$a_t = r\alpha. \tag{8.9}$$

Applying Equation 8.8b to an object set to rotate with a constant angular acceleration and an initial angular velocity ω_o gives its final angular velocity ω after a time interval, t, as

$$\omega = \omega_o + \alpha t. \tag{8.10}$$

This is the first equation of rotational motion that is analogous to the first equation of one-dimensional motion (see Equation 2.9). Defining the initial angular position of the object as θ_o and its final angular position as θ, the analogy between one-dimensional linear motion and angular motion enables one to express the remaining angular equations of motion as

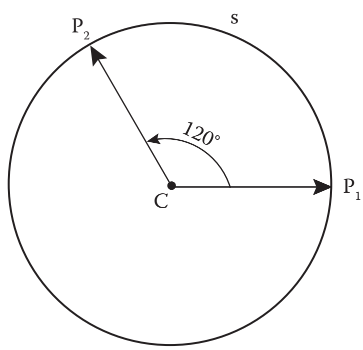
$$\theta - \theta_o = \bar{\omega} t, \tag{8.11}$$

$$\theta - \theta_o = \omega_o t + \frac{1}{2}\alpha t^2, \tag{8.12}$$

$$\omega^2 = \omega_o^2 + 2\alpha(\theta - \theta_o). \tag{8.13}$$

EXAMPLE 8.1

For the circle (r = 15.0 cm) in the figure below, calculate the arc length, s, that subtends an angle $\theta = 120°$ at the circle's center.



Solution

Using Equation 8.2a, θ in radians becomes

$$\theta_{rad} = \left(\frac{2\pi rad}{360_{deg}}\right)\theta_{deg} = \left(\frac{2\pi rad}{360°}\right)(120°) = 2.1 rad.$$

Using Equation 8.1, the arc length is

$$s = r\,\theta = (15.0 \text{ cm})(2.1 \text{ rad}) = 31 \text{ cm}.$$

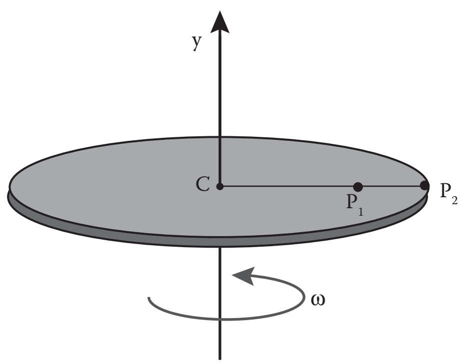
Analysis

Calculating the circumference C of the circle ($C = 2\pi r$), rounded off to two significant figures, gives a value of 94 cm. As the angle subtended by the arc s is 120°, one-third of the 360° full rotation, the length of the arc s ought to be one-third of the circumference. A slight difference from one-third is due to rounding off these values.

Example 8.2

A thin disk of 22.0 cm radius (the figure below) is rotating with a constant angular velocity $\omega = 33$ rev/min. Determine the linear velocity of

a. A point P_1 on the disk surface 14.0 cm from the center.
b. A point P_2 on the rim of the disk.



Solution

a. Upon converting ω from revolutions per minute to radians per second, it becomes

$$\omega = \left(\frac{33.0 \text{ rev}(2\pi \text{rad}/1.0 \text{ rev})}{1.0 \text{ min}(60.0 \text{ s}/1.0 \text{ min})}\right) = 3.5 \text{ rad/s}.$$

Using Equation 8.5 for point P_1 gives

$$v = r\omega = (14.0 \text{ cm})(3.5 \text{ rad/s}) = 49 \text{ cm/s}.$$

b. Again, using Equation 8.5 for point P_2 gives

$$v = (22.0 \text{ cm})(3.5 \text{ rad/s}) = 76 \text{ cm/s}.$$

Analysis

The point in part (b) has a higher linear velocity than its value in part (a) because its distance from the center is larger. The disk rotates with the same angular velocity, but different points on its surface have different linear velocities.

EXAMPLE 8.3

Consider a music record rotating at a rate of 78.0 rev/min. As the power is turned off, the record stops in 5.0 s. Determine

a. The angular acceleration of the record
b. The angle through which the record turns before it stops

Solution

a. Converting from revolutions per minute to radians per second, the initial angular velocity ω_o becomes

$$\omega_o = \frac{78.0\ \text{rev}(2\pi/1.0\ \text{rev})}{1.0\ \text{min}(60.0\ \text{s}/1.0\ \text{min})} = 8.2\,\text{rad/s}.$$

From Equation 8.10

$$\omega = \omega_o + \alpha t.$$

Thus,

$$0 = 8.2\ \text{rad/s} + \alpha(5.0\ \text{s}).$$

This gives

$$\alpha = -1.64\ \text{rad/s}^2.$$

b. From Equation 8.12

$$\theta - \theta_o = \omega_o t + \frac{1}{2}\alpha t^2.$$

Setting the starting angular position θ_o to 0, then

$$\theta = (8.2\ \text{rad/s})(5.0\text{s}) + \frac{1}{2}(-1.64\ \text{rad/s}^2)(5.0\ \text{s})^2 = 21\,\text{rad}.$$

Analysis

As noted, since the record is slowing down, the tangential acceleration is negative. One should note that there is also a centripetal acceleration experienced by this very point (see Chapter 3).

8.3 TORQUE

A torque can be defined by considering an object of mass m, attached to a solid rod, assumed to be of negligible mass, pivoted to rotate about an origin O (Figure 8.2). A force **F**, located at position vector **r** and making an angle φ with it and acting on the object, will rotate it about O. This is equivalent to saying that there is a torque, τ, about O acting on the object. The general definition of the magnitude of this torque is

$$\tau = rF \sin \varphi. \tag{8.14}$$

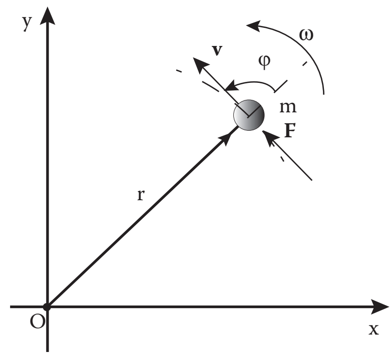


FIGURE 8.2 A torque on an object of mass m and position vector **r** is created by a force F acting on it along a direction that makes an angle φ with **r**.

If $\varphi = 90°$, that is, when the applied force is perpendicular to r, $\sin\varphi = 1$, and the acting torque becomes

$$\tau = Fr_{\perp}. \tag{8.15}$$

Equation 8.15 states that the magnitude of the torque is equal to the applied force times its perpendicular distance from the axis of rotation. As for the direction of rotation, it is in the plane that combines the arm r and force F. For a viewer facing the plane, the rotation would be clockwise or counterclockwise.

Using Newton's second law, substitution for F in Equation 8.15 gives

$$\tau = rma_t.$$

Substituting for a_t from Equation 8.9 in the above equation gives

$$\tau = mr^2\alpha \tag{8.16}$$

or

$$\tau = I\alpha, \tag{8.17}$$

where I, called the moment of inertia of the object about O, is given by

$$I = mr^2. \tag{8.18}$$

In Equation 8.18, r is the distance between the object and the axis of rotation passing through O perpendicular to r. Comparing Equation 8.17 with Newton's second law, $F = ma$, a distinct parallel between them is noted. As also can be noticed from Equations 8.17 and 8.18, both the moment of inertia and the torque depend on the location of the axis about which the object is rotating. In case several forces act on the object, the net torque is the vector sum of the torques acting on it. Hence, Equation 8.17 becomes

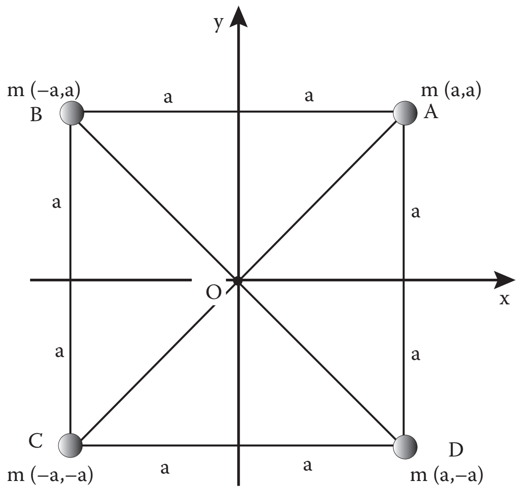
$$\tau_{net} = I\alpha. \tag{8.19}$$

In this treatment, forces applied to an object will be limited to those acting on it in one plane. As a matter of convention, torques that create rotations in a clockwise direction are considered negative and those that create rotations in a counterclockwise direction are considered positive.

EXAMPLE 8.4

A rigid body consists of four equal masses fastened in the xy plane (plane of the page, the figure below), m at (a, a), m at (–a, a), m at (–a, –a), and m at (a, –a). In terms of m and a, determine the following:

a. The moment of inertia about the x axis
b. The moment of inertia about the y axis
c. The moment of inertia about the z axis



SOLUTION

a. From Equation 8.18, the moments of inertia of the four objects about the x axis are, respectively,

$$I_1 = m\,(a^2),\ I_2 = m\,(a^2),\ I_3 = m\,(a^2),\ I_4 = m\,(a^2).$$

Thus, the total moment of inertia of the four objects is

$$(I_{tot})_x = I_1 + I_2 + I_3 + I_4 = 4\ ma^2.$$

b. Again, from Equation 8.18, the moments of inertia of the four objects about the y axis are, respectively,

$$I_1 = m\,(a^2),\ I_2 = m\,(a^2),\ I_3 = m\,(a^2),\ I_4 = m\,(a^2).$$

Thus, the total moment of inertia of the four objects is

$$(I_{tot})_y = I_1 + I_2 + I_3 + I_4 = 4\ ma^2.$$

c. As for the z axis, each object is distant from the origin O by a distance r equal to

$$r = \sqrt{(a^2 + a^2)} = \sqrt{2a^2} = a\sqrt{2}.$$

Thus, the moment of inertia of each object about the z axis, that is, about O, is

$$I = mr^2 = m(\sqrt{2}a)^2 = 2ma^2.$$

As the objects are equidistant from the origin, the moment of inertia of the whole system is then four times the moment of inertia of each. That is,

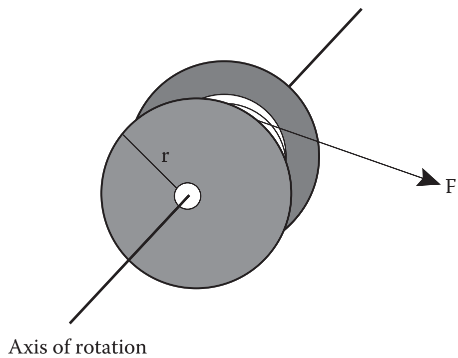
$$(I_{tot})_z = 4\,I = 8\,ma^2.$$

Analysis

1. Again, the symmetry of the system with respect to the x and y axes is why the moments of inertia about the x and y axes are equal.
2. The moment of inertia of the system about O is the sum of the moments of inertia about the x and y axes and, hence, is equal to four times the moment of inertia of any of the four objects about O.

EXAMPLE 8.5

The figure below shows a cylinder of radius r = 0.45 m whose moment of inertia about its fixed axis of rotation is $I_c = 14.4$ kg m^2. A constant force F = 55.0 N is applied to a cable wrapped around the cylinder's periphery. If it started from rest, determine its angular velocity after it has rotated 22 revolutions.



Solution

The torque acting on the cylinder is

$$\tau = rF$$
$$= (0.45\text{ m})(55.0\text{ N})$$
$$= 24.8\text{ Nm}.$$

From Equation 8.17, the torque τ acting on the cylinder is

$$\tau = I\,\alpha.$$

Thus, substituting for τ and I, the above equation becomes

$$24.8\text{ Nm} = (14.4)\,\alpha,$$

which gives

$$\alpha = 1.72\text{ rad/s}^2.$$

From the angular equation of motion (Equation 8.13),

$$\omega^2 = \omega_o^2 + 2\alpha(\theta - \theta_o).$$

Substituting for $\omega_o = 0$ and $\theta_o = 0$, the above equation becomes

$$\omega^2 = 0 + 2(1.72\ \text{rad/s}^2)(22\text{rev}\ (2\pi\ \text{rad/rev})).$$

Thus,

$$\omega = 21.8\ \text{rad/s}.$$

Exercise

1. Determine the tangential velocity of a point on the periphery after the cylinder has rotated 22 revolutions (see Equation 8.7).
2. Find how much time the cylinder took to complete the 22 revolutions (see Equation 8.10).

Analysis

The connection between the rotational motion and the linear motion is again observed in this exercise.

EXAMPLE 8.6

A thin uniform rectangular slab of negligible mass, 1.00 m long, is pivoted at its center to rotate freely in a vertical plane as shown in the figure below. Two masses, $m_1 = 0.25$ kg and $m_2 = 0.50$ kg, are firmly secured to the slab at 0.80 m and 0.30 m, respectively, from its left end. Calculate

a. The torque of each mass about the center of mass of the slab
b. The net torque on this slab about its center of mass

Solution

a. From the given geometry (the figure below), the two masses m_1 and m_2 will be at 30.0 cm and 20.0 cm, respectively, from the axis of rotation. The force acting on the right side of the slab will be the weight m_1g. Thus,

$$F_1 = m_1g = (0.25\ \text{kg})(9.80\ \text{m/s}^2) = 2.5\ \text{N}.$$

Similarly, the force acting on the left side will be the weight m_2g. That is,

$$F_2 = m_2g = (0.50\ \text{kg})(9.80\ \text{m/s}^2) = 4.9\ \text{N}.$$

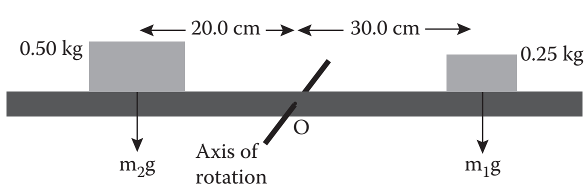
Thus,

$$\tau_1 = r_1\, F_1 \sin\varphi = (0.30\ \text{m})\ (2.45\ \text{N}) \sin 90.0° = 0.74\ \text{Nm, clockwise } (-)$$

and

$$\tau_2 = r_2\, F_2 \sin\varphi = (0.20\ \text{m})\ (4.90\ \text{N}) \sin 90.0° = 0.98\ \text{Nm, counterclockwise } (+).$$

b. The net torque acting on the slab is $\tau = \tau_2 - \tau_1 = 0.24$ Nm counterclockwise (+).



Analysis

Notice that all forces are in one plane, the vertical plane, and each mass will be moving in a rotational motion.

Exercise

Assuming that the system, described in the above example as being initially at rest, was observed for 22 s, find

1. Its angular acceleration at the end of this period
2. Its angular velocity at the end of this period

8.4 RIGID BODY

The previous discussion addressed single objects, each treated as a point particle of mass m rotating about a point, O. The treatment could be generalized to a system that consists of discrete objects attached firmly to each other via solid rods whose masses are assumed negligible (Figure 8.3). The treatment also addresses systems that are finite continuous distributions of mass (Figure 8.4). Each of these systems is identified as a rigid body.

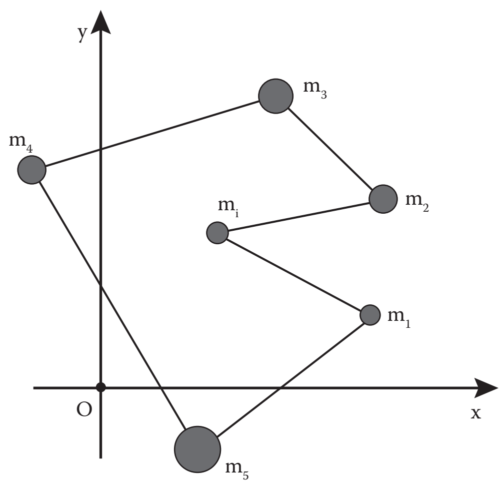


FIGURE 8.3 A rigid body that consists of discrete objects attached firmly to each other via solid rods of negligible masses.

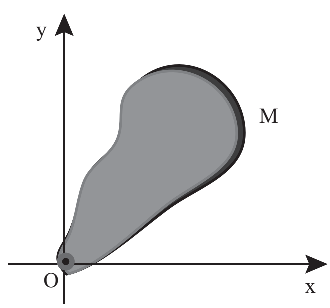


FIGURE 8.4 A rigid body that consists of a finite continuous distribution of mass.

8.5 MOMENT OF INERTIA

For a collection of discrete of particles, every object is treated as a member of a system whose total moment of inertia about an axis of rotation is the sum of the moments of inertia, I_1, I_2, I_3, …, I_i, …, I_n of all objects about the same axis. Thus for objects, m_1, m_2, m_3, …, m_i, …, m_n (Figure 8.5), the system's moment of inertia about an axis through the origin O perpendicular to the xy plane is

$$I = I_1 + I_2 + I_3 + \cdots + I_i + \cdots + I_{n.}$$

That is,

$$I = m_1 r_1^2 + m_2 r_2^2 + m_3 r_3^2 + \cdots + m_i r_i^2 + \cdots + m_n r_n^2, \tag{8.20}$$

where r_1, r_2, … are the distances, respectively, of the particles m_1, m_2, … from the origin O.

For a rigid body of a continuous mass distribution, the determination of its moment of inertia about an axis of rotation is more involved and requires the use of calculus. Therefore, without going through derivations, values of several chosen rigid body geometries are listed in Table 8.1. Notice the difference in the value of the moment of inertia of any of these when the axis of rotation is changed from one location to another. As this chapter is noncalculus based, examples and problems that require the calculation of the moment of inertia will be limited for discrete systems only.

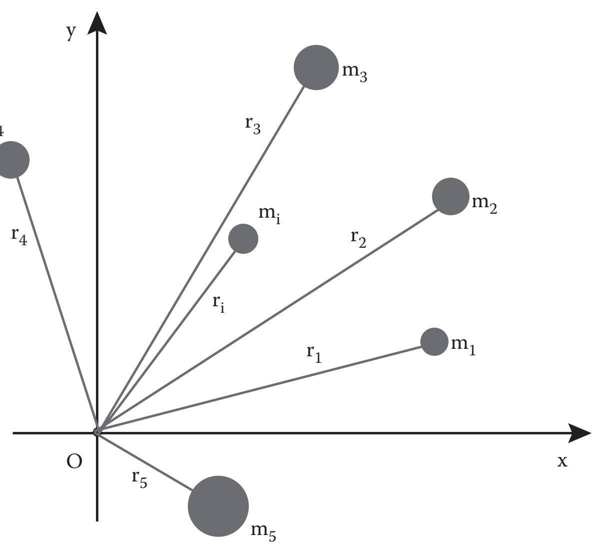
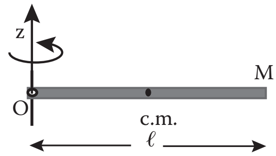


FIGURE 8.5 A system of discrete particles. The system's moment of inertia about an axis of rotation is the sum of the moments of inertia of all particles about the axis.

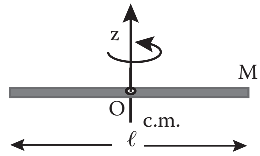
8.6 PARALLEL AXIS THEOREM

The moment of inertia of a rigid body about its center of mass has a special relevance because of its role in the analysis of its motion as will be noticed in Section 8.10. The moment of inertia of the rigid body (mass M) depicted in Figure 8.6, about an axis OZ′ through O, parallel to the axis CZ

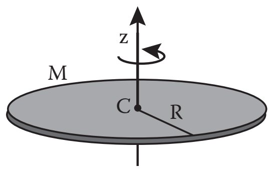
TABLE 8.1
Moment of Inertia for Selective Configurations of Rigid Bodies



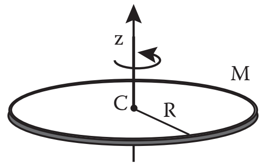
Thin rod of mass M and length ℓ about an axis through one end: $I = (1/3)\, M\ell^2$



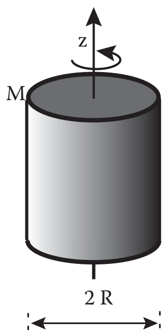
Thin rod of mass M and length ℓ about an axis through c.m. $I = (1/12)\, M\ell^2$



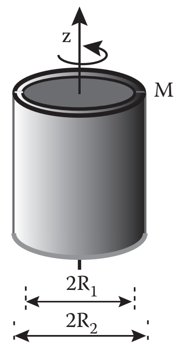
Disc of mass M and radius R about z: $I = (1/2)MR^2$



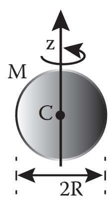
Hoop of mass M and radius R about z: $I = (1/2)MR^2$



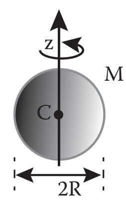
Solid cylinder of mass M and radii R about the central axis of symmetry: $I = (1/2)M\,R^2$



Thin cylindrical shell of mass M and radii R_1 and R_2 about the central axis of symmetry: $I = (1/2)M\,[(R_1)^2 + (R_2)^2]$



Thin spherical shell of mass M and radius R about any of its diameters: $I = (2/3)MR^2$



Solid sphere of mass M and radius R about any of its diameters: $I = (2/5)MR^2$

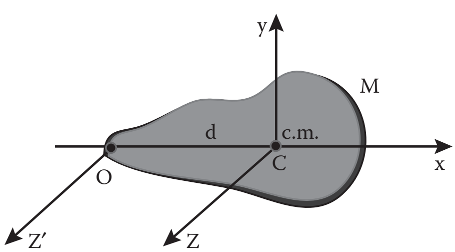


FIGURE 8.6 A rigid body with an axis of rotation CZ passing through its center of mass C. OZ′ is another axis of rotation parallel to CZ.

through the center of mass can be calculated by the parallel axis theorem which is expressed by the equation

$$I_o = I_{c.m.} + Md^2, \tag{8.21}$$

where d is the perpendicular distance between CZ and OZ′.

The parallel axis theorem is particularly useful when a uniform rigid body has both rotational and linear translational motions, as is the case of a cylinder or a sphere rolling on a straight flat surface or down an inclined plane.

8.7 EQUILIBRIUM OF A RIGID BODY

An object is in equilibrium when both its linear and angular accelerations are zero. This is possible only if the net force and net torque acting on the rigid body are zero. Thus,

$$\mathbf{F}_{net} = 0 \tag{8.22a}$$

and

$$\boldsymbol{\tau}_{net} = 0. \tag{8.22b}$$

Although each of the above conditions is necessary for a rigid body to be in equilibrium, neither of them in the absence of the other is sufficient. As an example, a uniform meter stick on which two forces, equal in magnitude and opposite in direction act at its two ends (Figure 8.7), cannot be in equilibrium, because the two forces create a torque that rotates the meter stick. Since $\mathbf{F}_{net}$ is a vector, condition (8.22a) implies that all components of the net force acting on the body should be zero for it to be in equilibrium. That is, for forces acting in an xy plane

$$F_{net,x} = 0, F_{net,y} = 0. \tag{8.23}$$

The same argument applies to Equation 8.22b. The torque $\boldsymbol{\tau}_{net} = 0$ implies that torques about z all be zero for the rigid body to be in equilibrium. Thus,

$$\tau_{net,any\ point} = 0. \tag{8.24}$$

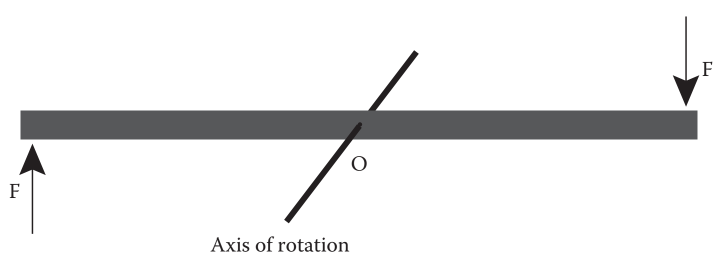


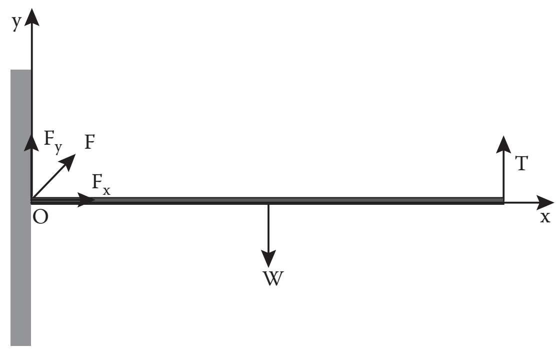
FIGURE 8.7 A uniform meter stick on which two equal and opposite forces are acting at its two ends that cannot be in equilibrium. The two forces create a torque about O.

EXAMPLE 8.7

A uniform metal bar of 1.20 m length and mass M = 1.40 kg is attached to a side vertical support at O. The bar is kept in equilibrium by pulling upward with a force T on it at its other end. Determine the tension and the force acting on the bar by the support.

Solution

The forces acting on the rod are the tension upward, the weight of the bar, W = mg, downward, and F of the support in a certain angle to be determined (the figure below).



Resolving **F** at O to two components F_x and F_y, and applying Equation 8.23 results in

$$F_x = 0 \tag{8.25}$$

and

$$F_y + T - W = 0.$$

Thus,

$$F_y + T - (1.40\ \text{kg})\ (9.80\ \text{m/s}^2) = 0.000$$

or

$$F_y + T - (13.7\ \text{N}) = 0.000. \tag{8.26}$$

From Equation 8.24, the net torque about point O must be zero. Thus,

$$F_y(0.00) + T(1.20\ \text{m}) - (1.40\ \text{kg})(9.80\ \text{m/s}^2)(0.600\ \text{m}) = 0.000.$$

This gives

$$T = 6.86\ \text{N}.$$

Substituting in Equation 8.26 yields

$$F_y + 6.86\ \text{N} - (13.7\ \text{N}) = 0.$$

Thus,

$$F_y = 6.8\ \text{N}.$$

Analysis

Notice that the torque was taken about the point of support at O. This ensures that the force of reaction F from the support creates no torque about that axis and, hence upon using Equation 8.22b, eliminates F, making the determination of T an easier task.

8.8 ANGULAR MOMENTUM

Referring to Figure 8.8, a point-like object at a distance r from the origin, rotating in the xy plane about the origin O with an angular velocity ω, has a linear velocity, $\mathbf{v} = \omega r$, and a linear momentum,

$$\mathbf{p} = m\mathbf{v}.$$

The object as treated for the moment is a point-like particle that has an angular momentum, **L**, of a magnitude defined as

$$L = rp \sin \varphi, \tag{8.27}$$

where φ is the angle between the position vector, **r**, and the momentum, **p**.

The angular momentum **L** is a vector whose direction is always perpendicular to the plane formed by the position vector **r** and the linear momentum **p**, that is, **L** is perpendicular to the plane of rotation.

If φ is 90°, then the magnitude of **L** is

$$L = rp = rmv. \tag{8.28}$$

Also since $v = \omega r$, then

$$L = mr^2 \omega$$

or

$$L = I\omega. \tag{8.29}$$

Analogous to linear motion ($P = mv$), the angular momentum of a rotating object, L, is equal to the product of the moment of inertia, I, and the angular velocity, ω. Equation 8.29 is a vector equation, and the angular velocity ω is treated as a vector with a direction perpendicular to the plane of rotation. This implies that for a point particle L and ω are parallel vectors. Thus, the vector form of Equation 8.29 is

$$\mathbf{L} = I\boldsymbol{\omega}. \tag{8.30}$$

Equation 8.30 applies to a rigid body as well. It shows that if the angular velocity of the rigid body ω is constant, its angular momentum is constant. As long as the axis of rotation of the rigid

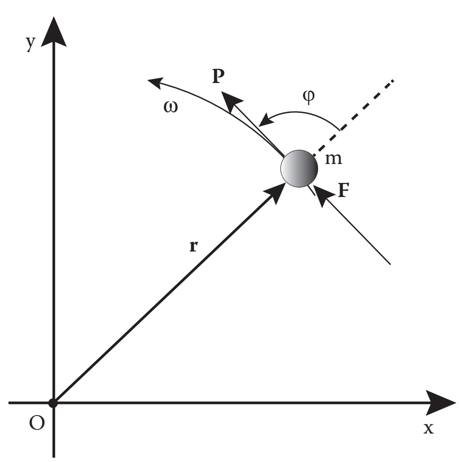


FIGURE 8.8 A point particle rotating in a circle has an angular momentum L.

body does not change, any change in its angular momentum must come from a change in its angular velocity, generating an angular acceleration. Finite changes in L and ω lead to

$$\frac{\Delta L}{\Delta t} = I\frac{\Delta w}{\Delta t}. \tag{8.31}$$

When these changes are considered over infinitesimal time intervals, the right hand side of Equation 8.31 becomes **I**⊠ (see Equation 8.17), which is the net torque created by a tangential force F acting on the rigid body. Therefore,

$$\frac{\Delta \mathbf{L}}{\Delta t} = \mathbf{I}\alpha = t_{net}. \tag{8.32}$$

The above three-term equation can be split into the following two equations:

$$\frac{\Delta \mathbf{L}}{\Delta t} = \mathbf{I}\alpha \tag{8.33a}$$

and

$$\frac{\Delta \mathbf{L}}{\Delta t} = t_{net}. \tag{8.33b}$$

The condition for proper use of Equations 8.33 is that both the moment of inertia I and the net torque have to be calculated with respect to the same axis.

Equations 8.33a and 8.33b lead to two conclusions. If $\boldsymbol{\tau}_{net} = 0$,

1. The angular velocity $\boldsymbol{\omega}$ remains constant, because $\alpha = 0$
2. The angular momentum **L** remains constant. This establishes the conservation of angular momentum of an object

EXAMPLE 8.8

A uniform rectangular metal slab of 0.90 m length and mass M = 0.48 kg is pivoted to rotate horizontally about its left end. The slab is set into motion with an initial angular velocity ω = 4.2 rad/s, slows down uniformly until it stops in 12 s. Determine

a. The angular acceleration of the slab
b. The slab's initial angular momentum
c. The initial torque that was needed to set the slab into motion

SOLUTION

a. The described slab is depicted in the following figure.
   From Equation 8.10

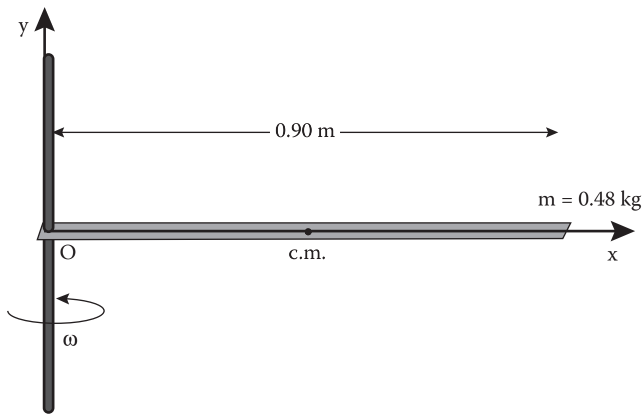
$$\omega = \omega_o + \alpha t.$$

Thus,

$$0 = 4.2 \text{ rad/s} + \alpha(12 \text{ s}).$$

This gives

$$\alpha = -0.35 \text{ rad/s}^2.$$



b. From Table 8.1, the moment of inertia of a rod about either end is

$$\mathbf{I} = \frac{1}{3}M\ell^2,$$

which, after substituting for M and ℓ, gives

$$I = \frac{1}{3}(0.48\text{kg})(0.90\text{m})^2 = 0.13 \text{ kg m}^2.$$

c. The slab's initial angular momentum is $L = I\omega$. Thus,

$$L = (0.13 \text{ kg m}^2)(4.2 \text{ rad/s}) = 0.54 \text{ kg m}^2\text{/s}.$$

From Equation 8.33b

$$\tau_{net} = \frac{\Delta L}{\Delta t}.$$

That is,

$$\tau_{net} = \frac{L_f - L_i}{\Delta t}.$$

Substituting for the initial and final angular momenta in the above equation gives the magnitude of the torque as

$$|\tau_i| = \left|\frac{0 - 0.54 \text{ kgm}^2/\text{s}}{12 \text{ s}}\right| = 0.045 \text{ kgm}^2/\text{s}^2.$$

There exist some subtleties, mainly regarding spatial directions, in Equations 8.33a and 8.33b. They arise because torque ($\boldsymbol{\tau}$) and angular momentum (**L**) are vector quantities and thus have spatial direction. The subtlety occurs because the vectors are perpendicular to the plane of the motion and not parallel to it as is the case with, say a velocity vector. As an example, consider an object revolving or rotating about an axis with angular momentum (**L**). The direction of (**L**) may be determined in the following fashion: (1) determine the imaginary plane of motion; (2) position yourself in space, such that the motion is in a clockwise sense as you view it; and (3) the direction in space at which you are looking at the clockwise rotation is the vector direction of **L**. Similarly, the argument applies for the torque direction. The following examples are intended to clarify these concepts.

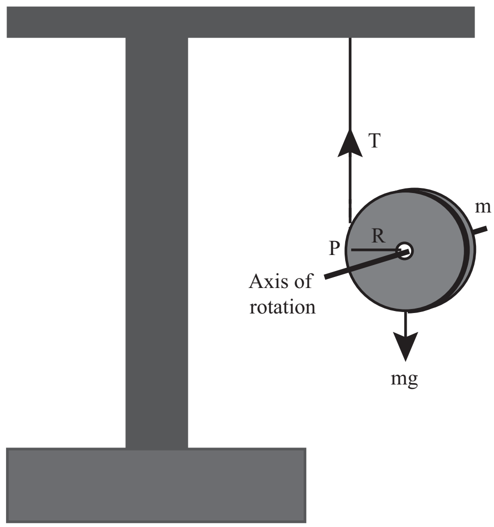
EXAMPLE 8.9

A string is wound in many turns around a circular disk of mass m and radius R. The "loose end" of the string is attached to the top of a heavy stand. The disk, released from rest, accelerates downward.

a. Determine the linear acceleration of the center of mass of the disk.
b. Determine the angular momentum of the disk, 10 s after release.

SOLUTION

a. The disk rotates about the point P, which instantaneously moves downward. The line of action of the tension in the string (T) passes through P and therefore contributes no torque about that point (the figure below).



The torque about point P is created by the disk weight (W) acting through its center of mass. Thus from Equations 8.14 and 8.19,

$$\tau_{net} = WR = I_p\alpha. \tag{8.34}$$

By the parallel axis theorem,

$$I_p = I_{cm} + mR^2.$$

The moment of inertia for the disk about its center of mass is $I_{cm} = (1/2)mR^2$. This reduces the above equation to the form

$$I_p = \frac{1}{2}mR^2 + mR^2 = \frac{3}{2}mR^2.$$

Also, as

$$WR = mgR \text{ and } a_{cm} = \alpha R,$$

Equation 8.34 becomes

$$mgR = I_p\alpha$$

$$= \left(\frac{3}{2}mR^2\right)\left(\frac{a_{cm}}{R}\right),$$

or

$$mgR = \left(\frac{3mR}{2}\right)a_{cm}.$$

Thus,

$$a_{cm} = \left(\frac{2}{3}\right)g.$$

b. Using Equation 8.32

$$\tau_{net} = \left(\frac{\Delta L}{\Delta t}\right) = I\alpha.$$

That is,

$$L_f - L_i = \Delta t I\alpha$$

or

$$L_f = L_i + (\Delta t)\tau_{net}$$

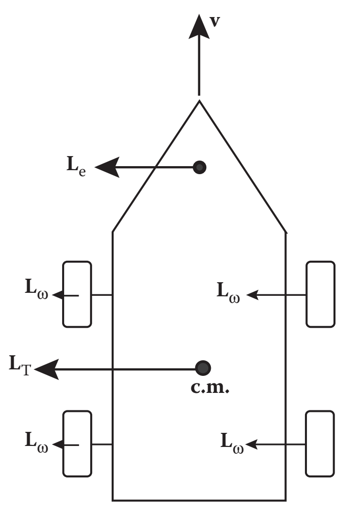
$$= 0.00 + (10.0\ s)(WR)$$

$$= 10.0\ mgR.$$

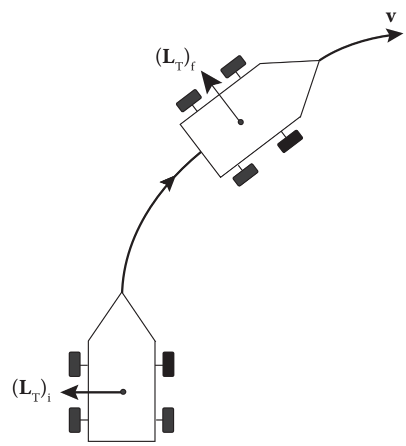
EXAMPLE 8.10

This example requires no numerical answer but investigates some consequence of Equations 8.33a and 8.33b. Consider an automobile with its engine mounted transversely. That is, if viewed from the right front fender, the running engine's crankshaft and flywheel rotate in a clockwise sense. By definition, this means that the angular momentum vector associated with the running engine ($\mathbf{L}_e$) is directed from the right to left of the car.

Similarly, when the auto is moving forward and is viewed from its right side, its wheels rotate clockwise and their angular momentum vectors ($\mathbf{L}_w$) are also directed from right to left. So, engine and wheel angular momenta are parallel and add directly to give a total angular momentum. Thus, $\mathbf{L}_T = \mathbf{L}_e + \mathbf{L}_w$ (the figure below).

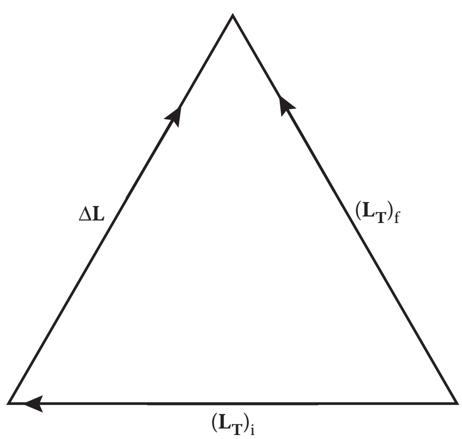


When a moving auto turns to the right or left, there is a change in the direction of its total angular momentum $\mathbf{L}_T$. This comes from a torque on the auto about an axis through its center of mass.



Depiction of a right turn at a constant speed in which the magnitude, but not the direction, of $\mathbf{L}_T$ remains constant.

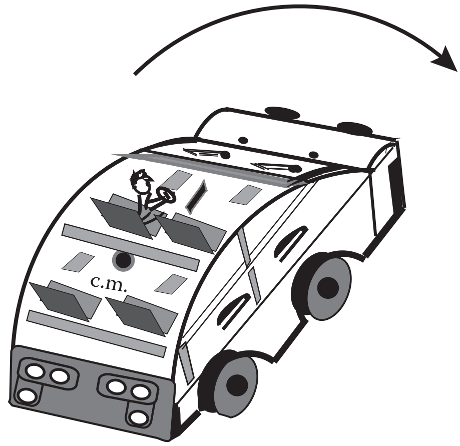
The figure above depicts a right turn at a constant speed in which the magnitude, but not the direction of $\mathbf{L}_{T,}$ remains constant. The figure below shows ($\Delta\mathbf{L}_T$). The net torque ($\boldsymbol{\tau}_{net}$) is in the same direction as $\Delta\mathbf{L}_T$, which, in this example, is along the direction of the car's motion.



The example shows ($\Delta\mathbf{L}_T$). The net torque ($\boldsymbol{\tau}_{net}$) is in the same direction as $\Delta\mathbf{L}$, which in this example, is along the direction of the car's motion.

The following figure depicts the auto from the rear, viewed along its direction of motion. Since ($\boldsymbol{\tau}_{net}$) is directed from rear to front, the apparent rotation is clockwise resulting in an additional force pushing down on the right side wheels. Correspondingly, ($\boldsymbol{\tau}_{net}$) would decrease the normal force on the left wheels. Thus, the right wheels "dig in" and are less prone to slide. An increase in the normal force results in an increase in the frictional force. Or, if the turn is made at a high speed, the auto has an increased possibility of rolling over.

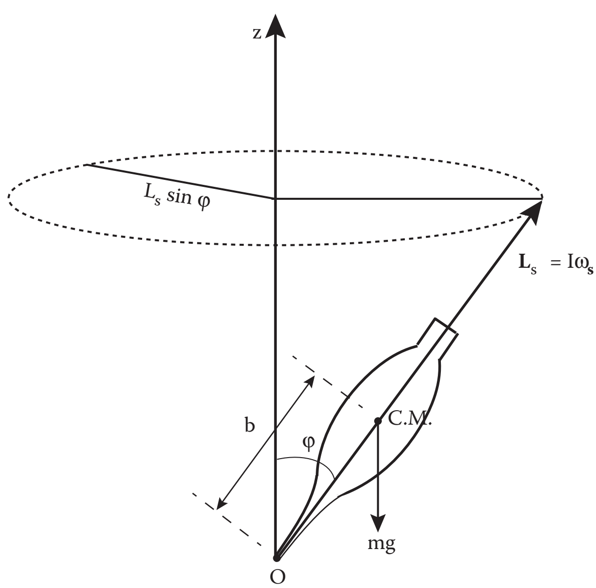
If the same auto makes a right turn, ($\Delta\mathbf{L}$) will be directed from front to rear as will ($\boldsymbol{\tau}_{net}$). The apparent rotation will be counterclockwise, as viewed from the rear. The left wheels now "dig in" and the normal force on the right wheels decreases. Thus, the car tends to "slide out" but not roll over on a left turn.



Depiction of the auto from the rear, viewed along its direction of motion.

A similar effect occurs with propeller-driven aircraft. When viewed from the front, the propellers rotate counterclockwise, so their original angular momentum is directed to the front of the plane. If a left turn is made, while the plane is in flight, the change in angular momentum ($\Delta\mathbf{L}$) is directed radially inward along the radius of curvature of the plane's flight path. The torque created tends to lift the rear and push down on the front of the plane.

An additional example of the effect of Equations 8.33a and 8.33b on flight is the helicopter. One purpose of the rear propeller is to apply a torque to prevent the helicopter body from rotating about the axis of the lift propeller. If this torque was not present, the helicopter body, immediately after lift-off, would start to rotate in a sense opposite to lift propeller, in order to conserve angular momentum. The rear propeller torque prevents conservation of angular momentum. An example of this situation is demonstrated in the figure below that depicts a rapidly spinning top at a given instant of time. The top has a mass m and its center of mass is located a distance (b) from the tip, along the axis of symmetry.



Depiction of a rapidly spinning top at a given instant of time. The top has a mass m and its center of mass is located a distance b from the tip, along the axis of symmetry.

The top has an angular momentum ($\mathbf{L}_s$), due to its spin, and experiences a torque, due to its weight, acting through the center of mass. The torque gives rise to a time rate of change of angular momentum (Equation 8.32), so the top precesses about its point of contact. The tip of the ($\mathbf{L}_s$) vector can be perceived as precessing about z and traversing the circular precessional path in the figure above. To find an expression for the precessional frequency (ω_p), consider the top view of the precession circle, as shown in the figure below.

Now

$$\omega_p = \frac{\Delta\varphi}{\Delta t} \tag{8.35a}$$

From the expression for an angle, specified in radians (the figure below),

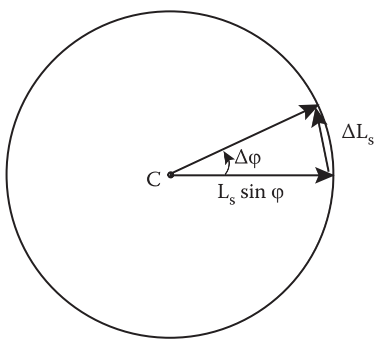
$$\Delta\varphi = \frac{\text{arc length}}{\text{radius}} = \frac{\Delta L_s}{L_s \sin\varphi}. \tag{8.35b}$$

Equation 8.35b into Equation 8.35a gives

$$\omega_p = \frac{1}{L_s \sin\varphi}\left(\frac{\Delta L_s}{\Delta t}\right). \tag{8.35c}$$

But from the previous figure,

$$\tau_{net} = \frac{\Delta L_s}{\Delta t} = W(b\sin\varphi). \tag{8.35d}$$



The top view of the precession circle of the spinning top depicted in the previous figure.

Equation 8.35d into Equation 8.35c yields

$$\omega_p = \frac{mg(b\sin\varphi)}{L_s(\sin\varphi)} = \frac{mgb}{L_s}. \tag{8.35e}$$

So the spinning top precesses, and the faster it spins, the larger is ($\mathbf{L}_s$) so the smaller is the rate of precession ω_p.

8.9 ROTATIONAL KINETIC ENERGY

The translational kinetic energy of a point-like object rotating in a circular path of radius r with a linear velocity v is

$$K = \frac{1}{2}mv^2. \tag{8.36a}$$

Substituting for $v = \omega r$ (see Equation 8.3), the above equation becomes

$$K = \frac{1}{2}mr^2\omega^2$$

or

$$K = \frac{1}{2}I\omega^2. \tag{8.36b}$$

Equation 8.36b demonstrates an analogy between the rotational kinetic energy of an object and its translational kinetic energy, K, that is equal to $(1/2)mv^2$, with the terms I and ω replacing the linear terms m and v, respectively. For a rigid body rotating about an axis through its center of mass, the kinetic energy of each of the particles constituting the rigid body is

$$(K)_i = \frac{1}{2}I_i\omega^2, \tag{8.36c}$$

where $I_i = m_i r_i^2$ is the moment of inertia of the ith particle about the axis of rotation.

The angular velocity ω about the axis of rotation is the same for all constituents of the rigid body, the rotational kinetic energy of the rigid body consisting of discrete particles about its axis of rotation is

$$(K_{body})_{rot} = \left(\sum_i \frac{1}{2}I_i\right)\omega^2, \tag{8.37}$$

while for a continuous distribution of mass, the above relation becomes

$$(K_{body})_{rot} = \frac{1}{2}I\omega^2. \tag{8.38}$$

8.10 A RIGID BODY IN TRANSLATIONAL AND ROTATIONAL MOTIONS

If a rigid body of mass M executes both translational and rotational motions, it can be shown that its total kinetic energy is always the sum of two contributions:

1. Rotational kinetic energy, which results from the rotation of the body about its center of mass. That kinetic energy is given by either Equation 8.37 or Equation 8.38, according to whether the body is a system of discrete particles (Figure 8.3) or a continuous mass distribution (Figure 8.4).
2. Translational kinetic energy, which is caused by the motion of the center of mass with a velocity V. This translational term has the velocity of the center of mass V. This kinetic energy is $(K)_{C.M} = (1/2)MV^2$. Thus, the total kinetic energy of the rigid body is

$$(K)_{rigid\ body} = \frac{1}{2}MV^2 + \frac{1}{2}I\omega^2. \tag{8.39}$$

8.11 TOTAL MECHANICAL ENERGY OF A RIGID BODY

The principle of the conservation of total mechanical energy for a rigid body in the absence of friction and any resistive dissipative forces applies to the dynamics of the rigid body in a manner similar to that for a point particle. Therefore, when only conservative forces act on the rigid body, the following equation

$$[(K) + (U)]_i = [(K) + (U)]_f = [(K) + (U)]_{\text{any arbitrary position}} \tag{8.40}$$

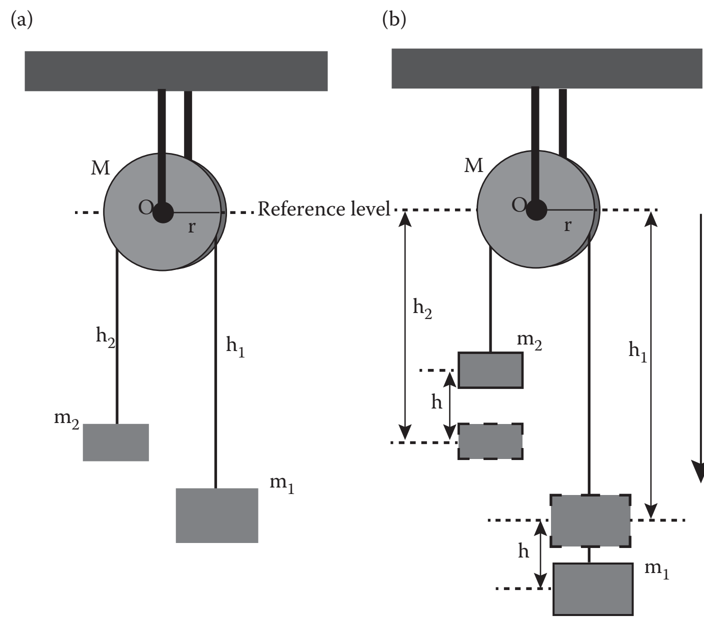
can always be applied. This will be demonstrated in a few examples.

EXAMPLE 8.11

Consider an Atwood machine that consists of a pulley of mass M and radius r that is free to rotate without friction on a frictionless axle through its center with two masses m_1 and m_2 ($m_1 = 2m_2$), hanging on the ends of an inextensible cord of length L around the pulley. The system is released from rest. Assuming that the cord does not slip, analyze the dynamics of the system and determine the velocity of both masses after m_2 has moved up (h_1 moved down) a distance h from its initial position.

SOLUTION

Let the two masses m_1 and m_2 be at heights h_1 and h_2 below the dashed reference level that passes through the pulley's center (the figure below, part (a)).



From Equation 8.40

$$[(K) + (U)]_i = [(K) + (U)]_f.$$

The initial kinetic energy of all elements (dashed blocks) of the system is zero. However, the initial potential energy of the system with respect to reference level is

$$[(U)_i]_{\text{system}} = [(U)_i]_{m1} + [(U)_i]_{m2} + [(U)_i]_{\text{Pulley}}$$

$$= (-m_1gh_1) + (-m_2gh_2) + 0.$$

Thus, the total mechanical energy E of the system is

$$E_i = 0 + (-m_1gh_1) + (-m_2gh_2) + 0. \tag{8.41}$$

To calculate the final kinetic energy of all the system, note that the speeds of the two masses are equal. Thus,

$$[K_f]_{system} = [K_f]_{m1} + [K_f]_{m2} + [K_f]_{Pulley}$$
$$= \frac{1}{2}m_1v^2 + \frac{1}{2}m_2v^2 + \frac{1}{2}I\omega^2.$$

The final potential energy of the system with respect to reference level (the figure above, part (b)) is

$$[U_f]_{system} = [U_i]_{m1} + [U_i]_{m2} + [U_i]_{Pulley}$$
$$= (-m_1g(h_1 + h)) + (-m_2g(h_2 - h)) + 0.$$

Thus, the total final mechanical energy is

$$E_f = \frac{1}{2}m_1v^2 + \frac{1}{2}m_2v^2 + \frac{1}{2}I\omega^2 + (-m_1g(h_1 + h)) + (-m_2g(h_2 - h)) + 0. \tag{8.42}$$

Equating the right hand sides of Equations 8.41 and 8.42 gives

$$(-m_1gh_1) + (-m_2gh_2) = \frac{1}{2}m_1v^2 + \frac{1}{2}m_2v^2 + \frac{1}{2}I\omega^2 + (-m_1g(h_1 + h)) + (-m_2g(h_2 - h)),$$

which reduces to

$$0 = \frac{1}{2}m_1v^2 + \frac{1}{2}m_2v^2 + \frac{1}{2}I\omega^2 - m_1gh + m_2gh.$$

Therefore,

$$\frac{1}{2}(m_1 + m_2)v^2 + \frac{1}{2}I\omega^2 = gh(m_1 - m_2).$$

Since the moment of inertia of the pulley, treated as a thin disk, is $1/2Mr^2$, and $v = \omega r$, then

$$\frac{1}{2}(m_1 + m_2)v^2 + \frac{1}{2}\left(\frac{1}{2}Mr^2\right)\left(\frac{v}{r}\right)^2 = gh(m_1 - m_2).$$

That is,

$$\frac{1}{2}(m_1 + m_2)v^2 + \left(\frac{1}{4}Mv^2\right) = gh(m_1 - m_2),$$

or

$$v = \sqrt{\frac{2gh(m_1 - m_2)}{(m_1 + m_2) + (1/2)M}}. \tag{8.43}$$

Analysis

1. Notice from the result in Equation 8.43 that if $m_2 = m_1$, the system would not move if it were initially released from rest.

2. If M were assumed of a negligible mass, then

$$v = \sqrt{\frac{2gh(m_1 - m_2)}{(m_1 + m_2)}}.$$

which is the same result that would be obtained for a simple Atwood machine.

EXAMPLE 8.12

Consider a solid sphere of mass m and radius R rolling down an inclined plane (the figure below). The incline makes an angle $\theta = 30°$ with the horizontal. If the sphere is released from rest, find its velocity after rolling without slipping a distance L along the incline.

Solution

From Equation 8.40, the total mechanical energy of the sphere is conserved. Thus,

$$[K + U]_i = [K + U]_f.$$

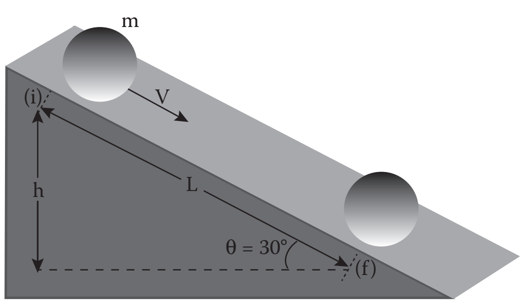
The initial kinetic energy of the sphere $(K)_i = 0$. However, its initial potential energy is

$$U = mgh = mgL \sin 30° = \frac{1}{2}mgL.$$

Thus, the initial total mechanical energy is

$$[K + U]_i = \frac{1}{2}mgL. \tag{8.44}$$

The final kinetic energy of the sphere $(K)_f$ is the sum of the translational kinetic energy of the center of mass and rotational kinetic energy about the center of mass.



That is, the final total mechanical energy

$$[K + U]_f = \frac{1}{2}mV^2 + \frac{1}{2}I\omega^2,$$

where $I = (2/5)mR^2$ is the moment of inertia of the sphere about its center of mass.

Thus, the final total mechanical energy becomes

$$[K + U]_f = \frac{1}{2}mV^2 + \frac{1}{2}\left(\frac{2}{5}mR^2\right)\omega^2. \tag{8.45}$$

Comparing Equations 8.44 and 8.45 gives

$$\frac{1}{2}mgL = \frac{1}{2}mV^2 + \frac{1}{2}\left(\frac{2}{5}mR^2\right)\omega^2.$$

As the sphere is rolling without slipping, $\omega = V/R$, and the above equation becomes

$$\frac{1}{2}mgL = \frac{1}{2}mV^2 + \frac{1}{2}\left(\frac{2}{5}mR^2\right)\left(\frac{V}{R}\right)^2.$$

or

$$\frac{1}{2}mgL = \frac{1}{2}mV^2 + \left(\frac{1}{5}mV^2\right).$$

That is,

$$\frac{1}{2}mgL = \left(\frac{7}{10}mV^2\right)$$

giving

$$V = \sqrt{\frac{5}{7}gL}. \quad (8.46)$$

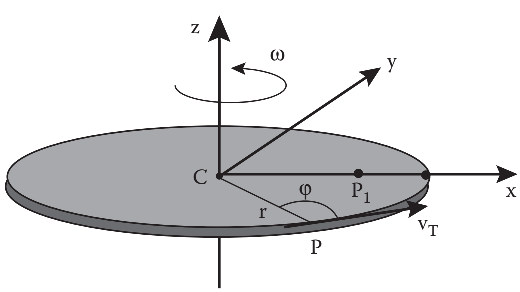
Analysis

1. The velocity of a point particle sliding down the same incline (see Example 4.9) is $V = \sqrt{2gh} = \sqrt{gL}$. The lower value of the velocity of the sphere's center of mass in Equation 8.46 resulted from the rotational motion.
2. As demonstrated above, Newton's second law was not used for determining the velocity of the center of mass of the sphere. However, the use of Newton's second law is a must if the acceleration of the center of mass is required.

PROBLEMS

8.1 An electrically powered disk in the xy plane (the figure below) of 78.0 cm radius is rotating about the z axis that passes through its center with an angular velocity $\omega = 18.0$ rad/s. When the motor is turned off, the disk comes to a stop in 24.0 s. For a point on the rim of the disk, determine

a. The linear velocity before the power is turned off
b. The angular acceleration during the disk slowing down
c. The linear (tangential) acceleration
d. The magnitude of the centripetal acceleration
e. The magnitude and direction of its total acceleration



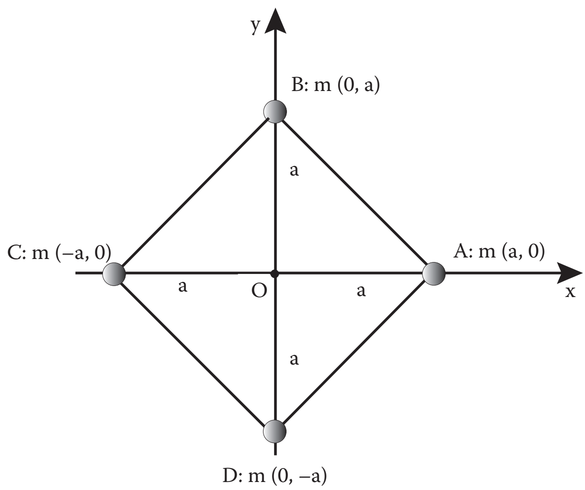
8.2 In the previous problem, if the moment of inertia of the disk is 0.035 kg m^2, what will be the magnitude and direction of its angular momentum just before the power is turned off?

8.3 Consider a grindstone, 16.0 cm diameter, that is electrically operated. As it is turned on, it rotates with a constant angular acceleration $\alpha = 33$ rad/s^2 before reaching a constant angular speed of 223 rad/s in a time interval t. If the grindstone's moment of inertia about its fixed axis of rotation is 1.28×10^{-3} kg m^2 and exhibits a frictional torque of 0.0022 kg m^2/s^2, determine

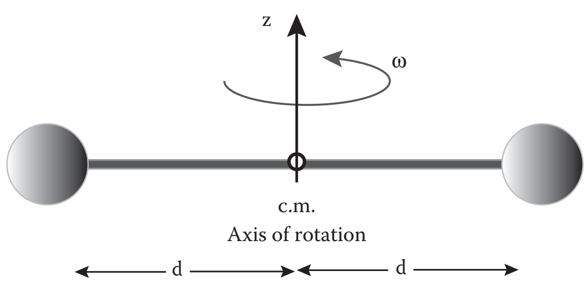
a. How many revolutions the grindstone has turned prior to picking up its constant speed
b. The time interval t it took the grindstone to pick up its constant speed
c. The electrically generated torque acting on the grindstone during the time interval t
d. The tangential acceleration of any point on the rim of the grindstone disk

8.4 A rigid body (the figure below) consists of four equal masses fastened in the xy plane (plane of the page) as follows: m at A (a, 0), m at B (0, a), m at C (–a, 0), and m at D (0, –a). If m = 0.400 kg and a = 3.00 m, determine the following:

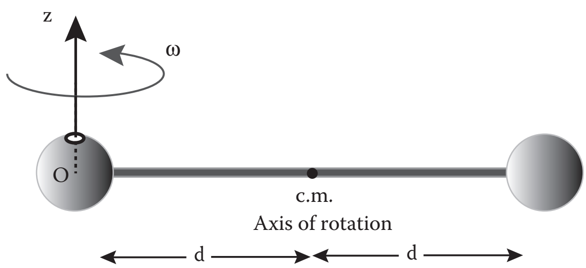
a. The moment of inertia about the x axis
b. The moment of inertia about the y axis
c. The moment of inertia about the z axis



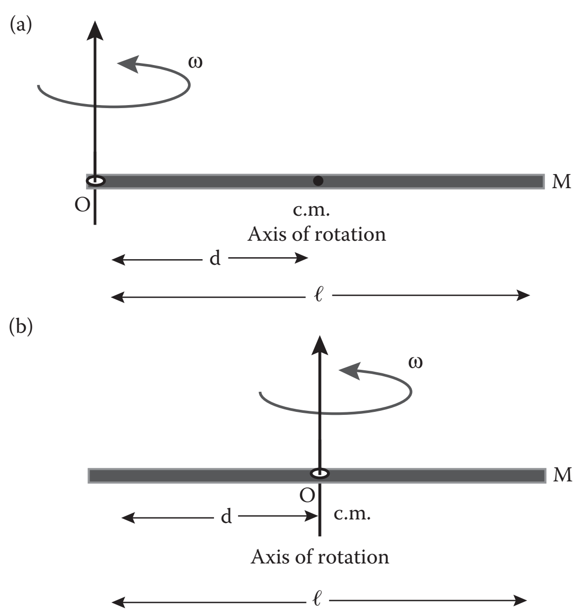
8.5 A simple dumbbell, consisting of two 5.20 kg spheres separated by a uniform 0.88 m rod of mass 2.20 kg, is rotating in a horizontal xy plane about its center of mass, c.m., with an angular velocity ω (the figure below). Determine dumbbell's moment of inertia about the rod's center of mass.



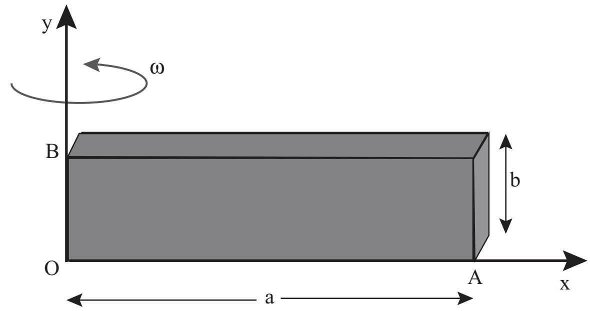
8.6 In the previous problem, calculate the moment of inertia about the z axis passing through the center of the left sphere at O (the figure below) in two ways:
a. By explicit calculations using the value for the moment of inertia of the rod from Table 8.1
b. By using the parallel axis theorem using your result in Problem 8.5



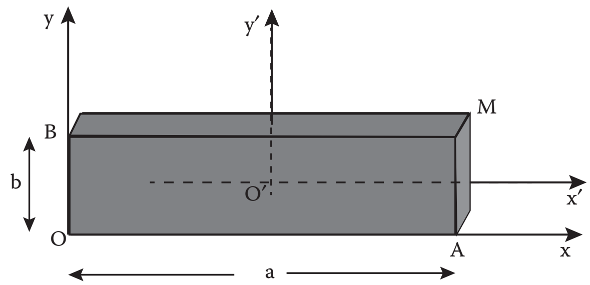
8.7 As stated in Table 8.1, for a rod of mass M and length ℓ, the moment of inertia about an axis y′ through one end perpendicular to the rod is $I = (1/3)M\ell^2$ (the figure below, part (a)), and the value about an axis $y_{c.m}$ through its center perpendicular to the rod is $I = (1/12)M\,\ell^2$ (part (b)). Show that these values satisfy the parallel axis theorem.



8.8 A thin uniform rectangular plate of mass m, length a, and width b is in the plane xy, where the x and y axes are directed along OA and OB, respectively (the following figure). The plate is hinged so that it could rotate freely about either the x or the y axis. Calculate
a. The plate's moment of inertia in terms of m, a, and b about the y axis
b. The plate's moment of inertia in terms of m, a, and b about the x axis

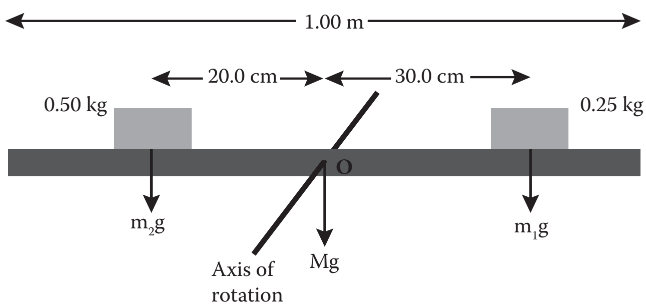


8.9 Use the results obtained in the previous problem to calculate the moments of inertia in terms of m, a, and b of the plate about the x′ and y′ axes that are parallel to the edges of the plates, passing through its center (the figure below).



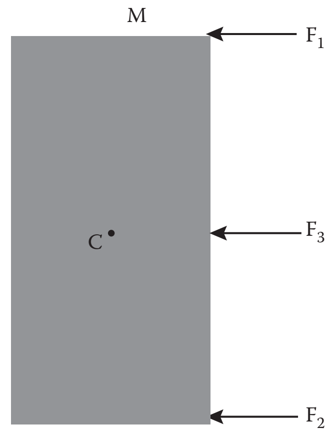
8.10 Consider a slab, mass M = 0.400 kg, on which two blocks are set at the positions shown in the figure below. Calculate

a. The moment of inertia of the system about O
b. The angular acceleration of the system

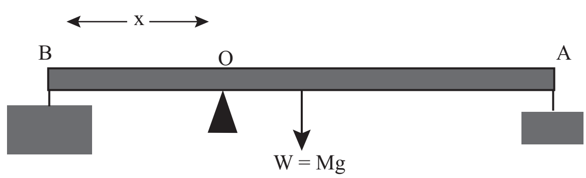


8.11 A flat uniform rectangular plate, 8.00 cm long, 4.00 cm wide, and mass M, is acted on by two force, F_1 = 11.0 N and F_2 = 14.0 N, at two of its corners, and F_3 = 77 N acting parallel to and midway between F_1 and F_2 (the following figure). If the plate is in the xy plane and is pivoted to rotate about a horizontal axis, the z axis, perpendicular to its plane through its center C,

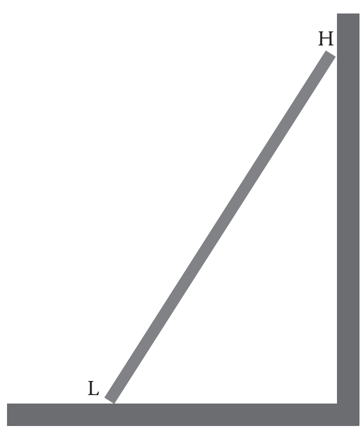
a. Determine the net torque acting on the plate.
b. Knowing that the plate's moment of inertia is 0.220 kg m^2, determine its angular acceleration.



8.12 The figure below shows a rod of uniform density, 1.00 m long and 11.0 N weight. A 12.0-N weight is hanging from the rod's end A, while a 28.0-N weight is hanging from the other end B. At what position, O, a distance x from B, should one place a fulcrum beneath the rod so that it would be in a stable equilibrium?

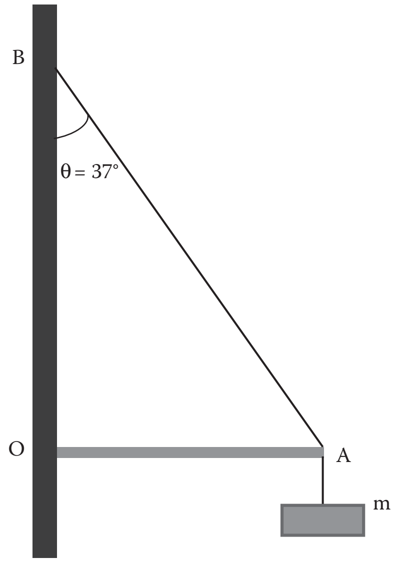


8.13 A uniform ladder of length 2.20 m and mass 6.00 kg is leaning against a frictionless wall with its lower end on the ground, 0.70 m away from the wall (the figure below). Determine the forces of reaction on the ladder from the wall and the ground.

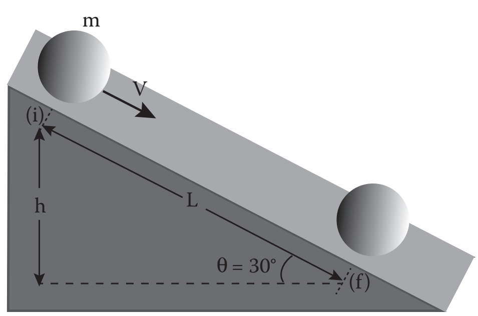


8.14 A mass of 14.0 kg is suspended from a uniform beam OA that is attached to a wall at O (the following figure). The beam, 1.00 m long and of 29.4 N weight, is held horizontally via a cable that is attached to the wall making an angle of 37.0° with it at point B. Determine

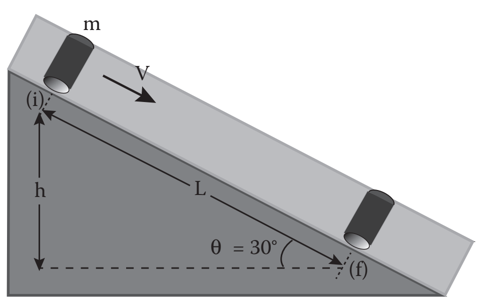
a. The tension in the cable
b. The forces of reaction at O and tension force of support along BA



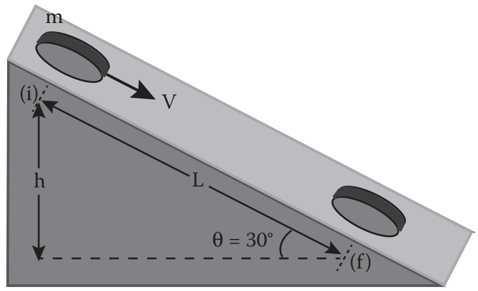
8.15 Consider the sphere, mass m and radius r (the figure below), that is rolling down an incline of length L = 1.0 m without slipping. Find
a. The acceleration of the sphere as it rolls down the incline
b. The time it takes the sphere to roll down the incline a distance L



8.16 Consider a solid cylinder of mass m and radius r rolling down an inclined plane (the figure below). The incline makes an angle θ = 30.0°. If the cylinder is released from rest, find
a. Its velocity in terms of g and L after rolling without slipping a distance L along the incline
b. The value of v in part (a) for L = 1.0 m



8.17 Consider a disk of mass m and radius r rolling down an inclined plane (the figure below). The incline makes an angle $\theta = 30.0°$ with the horizontal. If the disk is released from rest, find its velocity after rolling without slipping a distance $L = 1.0$ m along the incline.



9 Simple Harmonic Motion

Two kinds of motion, linear and rotational, have been studied so far. Another type, known as the simple harmonic motion (SHM), where the acceleration of the object is not constant is of special interest. The acceleration of an object executing SHM is proportional and opposite in direction to the displacement of the object from its equilibrium position. This is because the force acting on the moving object is proportional, but opposite to its displacement. The simple pendulum and the oscillation of a block, attached to an ideal spring, once displaced slightly from the equilibrium position are examples of SHM.

9.1 HOOKE'S LAW

Hooke's law demonstrates the response of linear elastic media when acted upon by a force. Springs are a good example. The situation is demonstrated in the following where a spring of a natural length ℓ (Figure 9.1a), attached to a rigid support at one end, has a mass m attached to its other end. The mass is pulled a distance x by a force F_{EXT} (Figure 9.1b). As long as the displacement is within the spring's elastic limit, the spring exhibits a self-created force F_{SP}, equal in magnitude but opposite in direction to the applied force, F_{EXT}. Once F_{EXT} is removed, the spring force becomes the only force acting on m, pulling it back toward its original position, O. That is why F_{SP} is called *the restoring force*. In all cases of a mass-spring system, the spring is assumed to be of a negligible mass.

Hooke's law established that the external force F_{EXT} needed to stretch a spring an amount x is

$$F_{EXT} = kx \tag{9.1}$$

and the spring force is

$$F_{SP} = -kx, \tag{9.2}$$

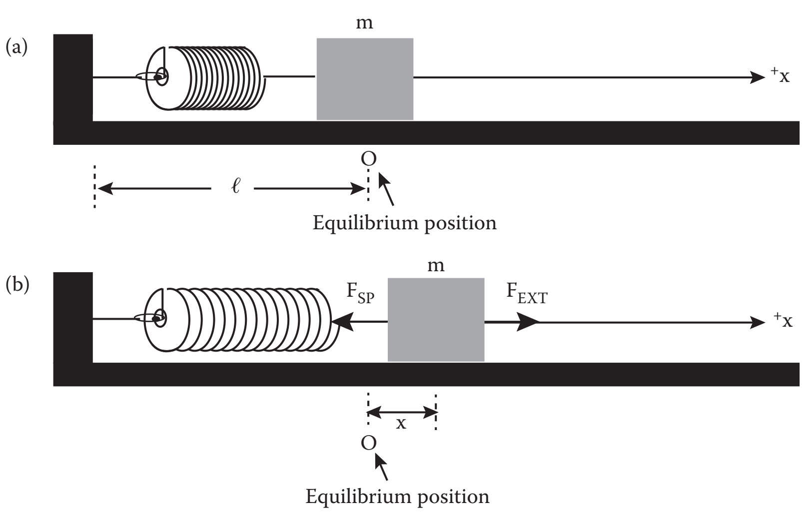


FIGURE 9.1 (a) Illustration of a horizontal spring of natural length ℓ attached to a rigid support on one end and mass m attached to its other end. (b) The mass is pulled a distance x from O by a force F_{EXT}.

k is a constant of proportionality called the spring constant, measured in N/m. The negative sign in Equation 9.2 indicates that the spring force is opposite to the displacement that is measured from the equilibrium position, O.

A similar case, a spring left to hang down from a solid support with a block of mass, m, attached to its free end is shown in Figure 9.2a. The position of m at O marks its equilibrium position. The figure shows the position of m after pulling on it to a new position, B (bottom position), by an external force, F_{EXT}, which, once removed, allows the mass to move upward. Subsequently, an oscillation of m up and down about the equilibrium position is initiated.

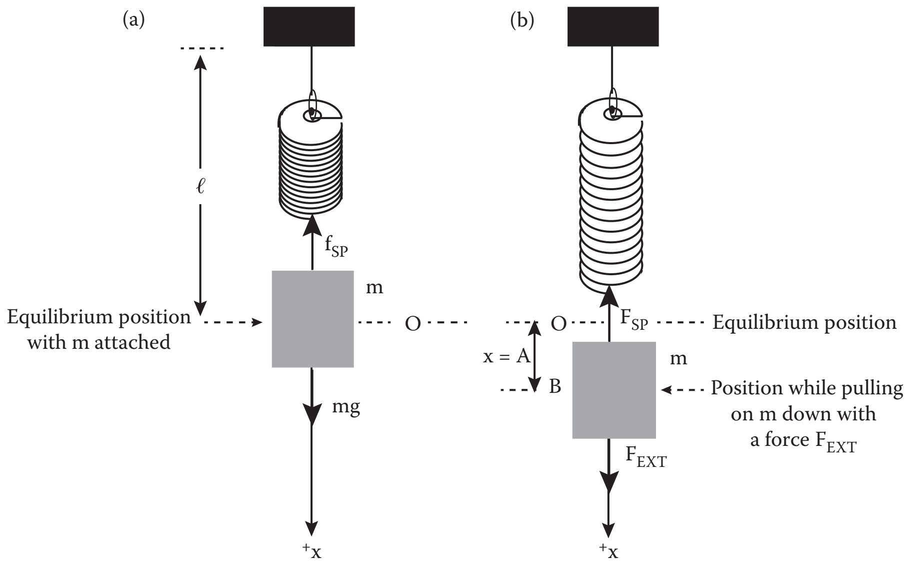


FIGURE 9.2 Illustration of (a) a vertical spring attached to a rigid support on one end and mass m attached to its other end. In this situation, $F_{SP} = mg$. (b) The mass is pulled a distance x = A below O by a force F_{EXT}. O is the equilibrium position of the mass.

EXAMPLE 9.1

A spring (the following figure, part (a)), of a natural length ℓ and spring constant k = 45.0 N/m, is laid out horizontally on a flat table with one end attached to a vertical side support and the other end attached to a block of mass m = 0.64 kg. Determine the force needed to pull the spring 14.0 cm away from its equilibrium position (part (b)).

Solution

From Equation 9.1,

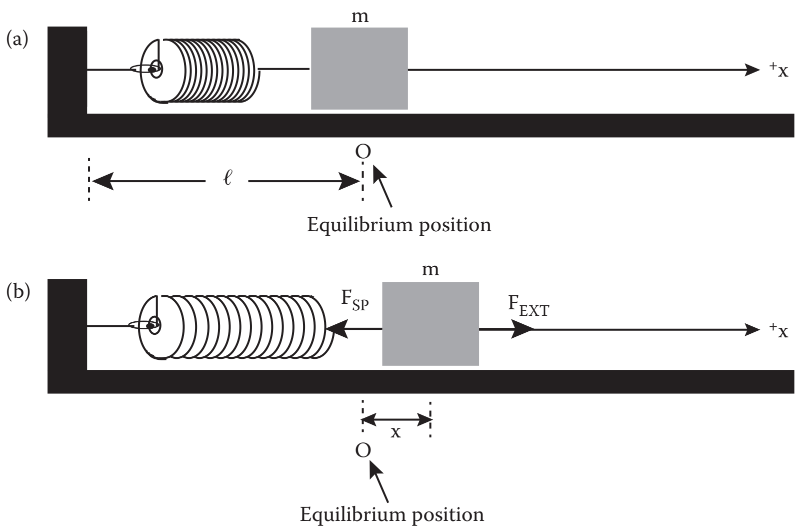
$$F_{EXT} = kx.$$

Thus,

$$F_{EXT} = (45.0\ \text{N/m})(0.14\ \text{m}) = 6.3\ \text{N}.$$

Analysis

As the spring force F_{SP} at any instant is equal to the external force applied to the spring when in equilibrium, then F_{SP} just prior to releasing the block will be 6.3 N.



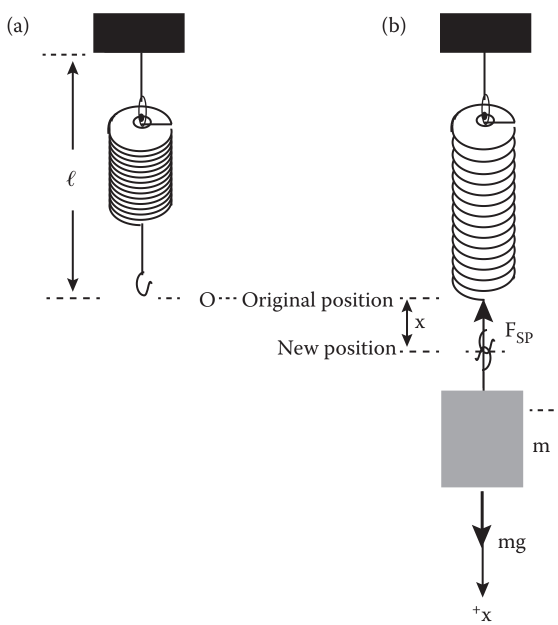
EXAMPLE 9.2

If a spring identical to that used in the previous example ($k = 45.0$ N/m) is hung from the ceiling (the figure below), find by how far down a block of 0.64 kg mass would stretch the spring beyond its natural length ℓ.

Solution

From Equation 9.1,

$$F_{EXT} = kx.$$



The external force in this case is the weight of the block. Thus,

$$mg = (k)(x)$$

or

$$(0.64 \text{ kg})(9.80 \text{ m/s}^2) = (45.0 \text{ N/m})(x),$$

which reduces to

$$6.3 \text{ N} = (45.0 \text{ N/m})(x).$$

Thus,

$$x = 0.14 \text{ m}.$$

Analysis

It can be observed that because the amounts of stretch of the spring in both cases were equal, the external force applied horizontally (6.3 N) had to be equal to that induced vertically down by the weight of the block (mg = 6.3 N).

9.2 POTENTIAL ENERGY OF A SPRING

As stated in Equation 9.1, the external force that acts on the spring, displacing it from equilibrium position O by an amount x, is

$$F_{EXT} = kx.$$

A plot of the force, F_{EXT}, versus the displacement, x, is shown in Figure 9.3. The work done by the external force is the shaded area in Figure 9.3. That is,

$$W_{EXT} = \frac{1}{2}kx^2. \tag{9.3}$$

This work is stored in the spring as potential energy U. That is, $U = W_{EXT}$. Hence,

$$U = \frac{1}{2}kx^2. \tag{9.4}$$

Equation 9.4 shows that the potential energy of the spring is positive for all values of x.

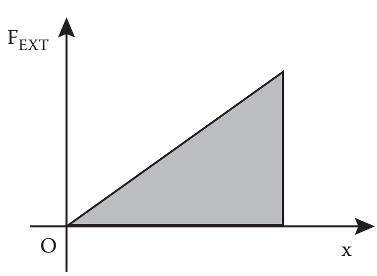
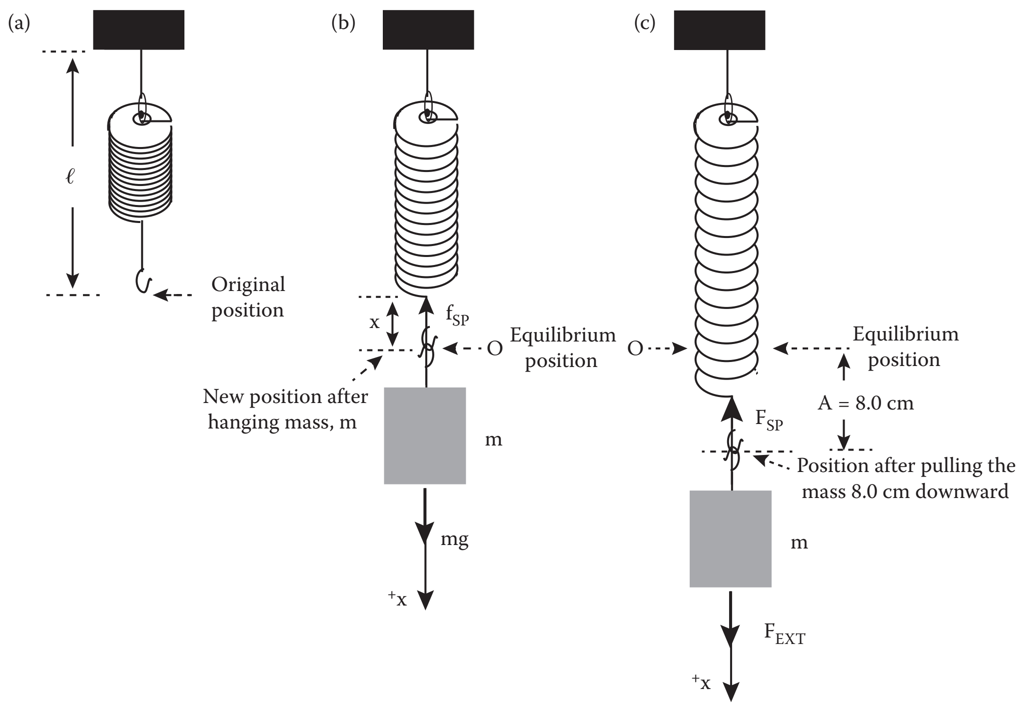


FIGURE 9.3 A plot of the force F_{EXT} versus the displacement, x, the shaded area represents the work done on the mass by the external force.

EXAMPLE 9.3

Use the information given in Example 9.2 and the figure below to determine

a. The work that gravity does on the spring (the figure below, part (a)) in stretching it 0.14 m (part (b))
b. The work done on the system to pull the block an additional 8.0 cm below its equilibrium position (part (c))
c. The total potential energy stored in the spring



Solution

a. In this part (the figure above, part (b)), the external force acting on the spring is the gravitational force, mg. As this force induces an extension of 14 cm, then the work it does on the spring is

$$W_{mg} = \frac{1}{2}(45.0 \text{ N/m})(0.14 \text{ m})^2 = 0.44 \text{ J}.$$

b. The additional 8.0 cm stretch is caused by an external force whose work would be

$$W_{EXT} = \frac{1}{2}(45.0 \text{ N/m})(0.08 \text{ m})^2 = 0.14 \text{ J}.$$

c. The total potential energy stored in the spring is equal to the total work W done on it. Thus,

$$W = W_{grav} + W_{EXT} = 0.44 \text{ J} + 0.14 \text{ J} = 0.58 \text{ J}.$$

Analysis

As can be observed, the potential energy in part (a) is due to the work done on the spring by the gravitational force, while in part (b) it is due to an external force pulling on the spring 8.0 cm further. In a horizontally laid out spring, the former stretch does not exist, and only the external force stretches the spring beyond equilibrium.

9.3 MASS-SPRING SYSTEM IN SHM

SHM is an oscillatory motion that is ideally periodic in space and time. In the absence of friction, this motion is executed by any object subjected to a net force proportional to and opposite to its displacement from equilibrium. As explained earlier, a mass-spring and simple pendulum are capable of executing SHM. This is illustrated in Figure 9.4.

In a horizontal mass-spring system (Figure 9.4a), the mass is manually pulled, stretching the spring and is then released (Figure 9.4b). The mass, acted on by the spring only, is pulled back toward its equilibrium position that it passes. The momentum acquired allows it to compress the spring further past its equilibrium position with gradually diminishing velocity to zero at L (Figure 9.4c). The spring, compressed to its maximum, exerts a force on m to the right, pushing it back

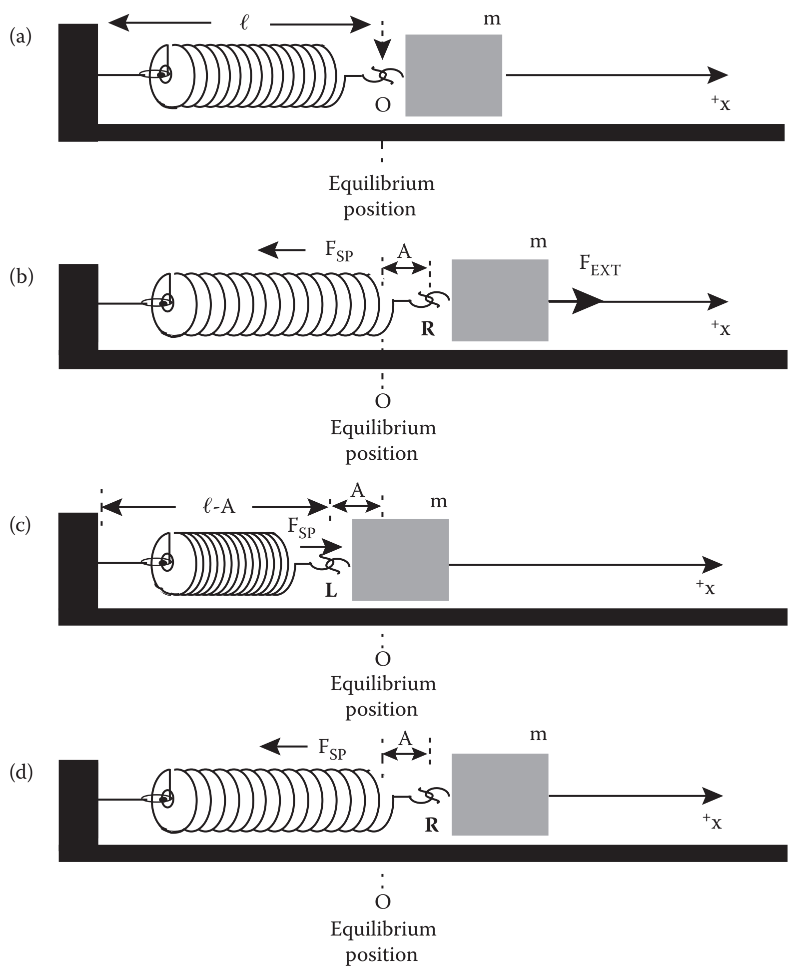


FIGURE 9.4 A mass spring system executing an SHM. (a) The spring of natural length ℓ is laid on a horizontal surface. (b) The spring is manually pulled, stretching it an amount A and is then released. (c) The mass, acted on by the spring, is pulled back toward its equilibrium position which then passes with gradually diminishing velocity to zero at L. The spring, compressed to its maximum, exerts a force on m to the right, pushing it back toward the equilibrium position, passing it and stretching the spring further to R as shown in (d).

toward the equilibrium position. Again, the mass gradually gaining momentum passes this position and stretches the spring to its maximum at R (Figure 9.4d), where one complete oscillation from R to L to R has been executed. This oscillation is also called a cycle, and as the system moves back and forth in sustained cycles, it is executing an SHM. The displacement from O to R through which m was pulled prior to releasing it is called the amplitude of the oscillation and is denoted by A.

9.3.1 Kinetic Energy and Total Mechanical Energy of Mass-Spring System in SHM

It has been established in Chapter 6 that the total mechanical energy E of an object acted upon by conservative forces is conserved. As the spring force is conservative, then at any position, x, the mass-spring system has a total mechanical energy

$$E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2 \tag{9.5}$$

that retains the same value. Hence, in Figure 9.4, equating the system's total energy at point R to its value at the origin O, gives

$$E_R = E_O,$$

or

$$\left(\frac{1}{2}mv^2\right)_R + \left(\frac{1}{2}kx^2\right)_R = \left(\frac{1}{2}mv^2\right)_o + \left(\frac{1}{2}kx^2\right)_o.$$

Let OR = A. As the block's velocity at R is zero, the above equation reduces to

$$0 + \frac{1}{2}kA^2 = \left(\frac{1}{2}mv^2\right)_o + 0.$$

Also, as m has its maximum velocity at O, the above equation becomes

$$\frac{1}{2}kA^2 = \frac{1}{2}mv_{max}^2 \tag{9.6}$$

or

$$v_{max} = \pm\sqrt{\frac{k}{m}}A = \pm\omega A, \tag{9.7}$$

where

$$\omega = \pm\sqrt{\frac{k}{m}} \tag{9.8}$$

is the angular frequency of the block. The plus or minus signs in Equation 9.8 signify the directions of the movement of the block as it goes through the equilibrium position. The positive sign for v_{max} indicates that the velocity of the block at the point x = 0 is moving toward the positive x axis chosen here to the right, and the negative sign indicates that the velocity of the block at the point $x = 0$ is moving toward the negative x axis. In each cycle of oscillation, m will be passing through O twice, one time on its way toward R and the other is on its way toward L.

Applying the energy conservation principle to this system one more time by equating the value of its total energy when m is at an arbitrary position x from O to its value when m is at R (OR = A) gives

$$\frac{1}{2}mv^2 + \frac{1}{2}kx^2 = \frac{1}{2}kA^2$$

or

$$v^2 + \frac{k}{m}x^2 = \frac{k}{m}A^2.$$

Since from Equation 9.8

$$\omega^2 = \frac{k}{m},$$

then

$$v^2 = \omega^2(A^2 - x^2)$$

and

$$v = \pm\omega\sqrt{(A^2 - x^2)}. \tag{9.9}$$

Equation 9.9 is a general expression that gives the velocity of m for any displacement, x. It can be observed that the special cases for which $x = 0$ gives $v_{max} = \pm\omega A$, and for $x = \pm A$, the velocity, $v = 0$, as expected.

The previous discussion applies as well to vertically oscillating springs (Figure 9.5). The oscillation of the mass about the equilibrium position O occurs between the lowest downward (D) and highest upward (H) positions.

9.3.2 Periodic Representation of Displacement in SHM

The property of SHM of an object whose displacement from equilibrium is ideally repetitive in space and time with an amplitude, A, allows the displacement to be describable by a sinusoidal, that is, sine or cosine, function. Considering a time-dependent cosine function, and referring to Figure 9.4,

$$x = A\cos(\omega t), \tag{9.10}$$

where ω is the angular frequency introduced earlier and is related to the linear frequency by

$$\omega = 2\pi f. \tag{9.11}$$

Here, f is called the linear frequency and is the number of oscillations per second, which is equal to the inverse of the period of the oscillation. That is,

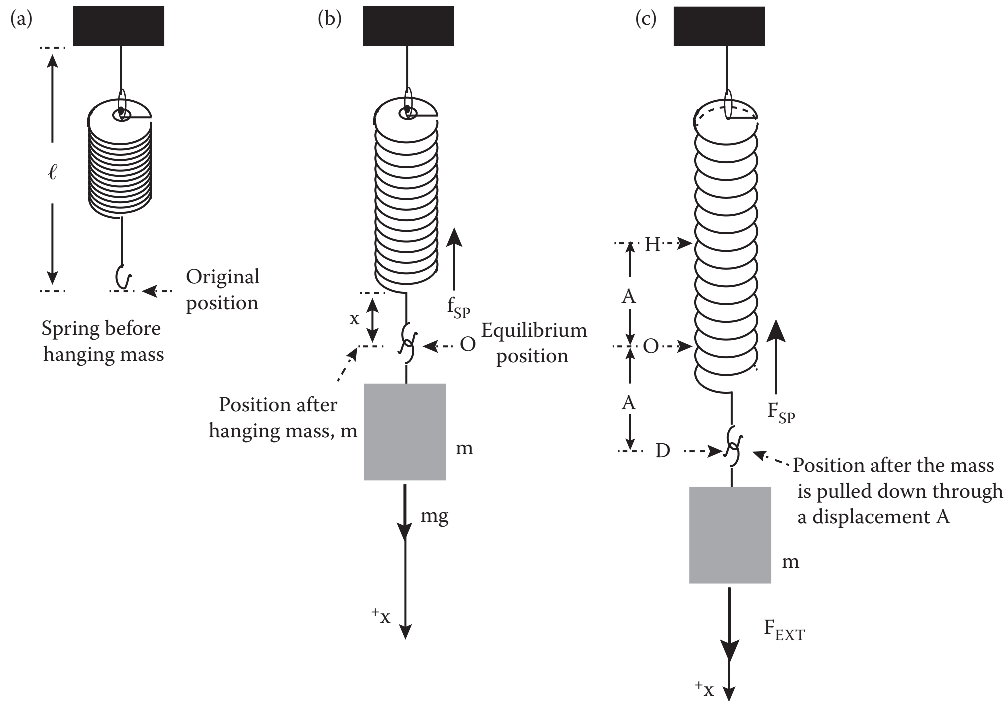


FIGURE 9.5 A vertical mass spring system executing an SHM about the equilibrium point O. D and H represent the lowest downward and highest upward positions, respectively, of the mass during oscillation.

$$f = \frac{1}{T}. \tag{9.12}$$

One full oscillation is known as a cycle. The description of a single oscillation as one cycle applies to any system executing an SHM. For a vertically oscillating mass-spring system, the oscillation would be up and down. This situation will be explained through several examples that are solved in the remaining sections of the chapter.

Equation 9.10 is a simplified version of x that describes the oscillation when it starts at x = A. A more general form is

$$x = A\cos(\omega t + \varphi), \tag{9.13}$$

where φ is called the initial phase angle or phase constant of the motion and can be determined from

$$\varphi = \cos^{-1}\left(\frac{x}{A}\right). \tag{9.14}$$

The following are several interesting phase constants that may mark various starting instants of the SHM:

1. x = A: φ = 0.0 rad. The observation starts when m is at the extreme right position, R (OR = A), in the horizontally oscillating spring or when it is at its lowest position, D, for the vertically oscillating spring.

2. $x = A/2$: $\varphi = (\pi/2)$ or $(3\pi/2)$ rad. The observation starts as m passes the equilibrium position, O, in both the horizontally and vertically oscillating springs. The two values of φ give the possible directions of motion of m toward the negative or positive directions of x, respectively.
3. $x = -A$: $\varphi = (\pi)$ rad. The observation starts when m is at the extreme left position, L, in the horizontally oscillating spring or when it is at its top most position H for the vertically oscillating spring.

The periodic motion of the mass in the mass-spring system expressed in Equation 9.13 extends to its velocity and acceleration as well. Without derivation, this motion is enhanced upon introducing the velocity in the form

$$v = -\omega A \sin(\omega t + \varphi) \tag{9.15}$$

and the acceleration in the form

$$a = -\omega^2 A \cos(\omega t + \varphi).$$

That is,

$$a = -\omega^2 x. \tag{9.16}$$

From Equations 9.8 and 9.11

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}}. \tag{9.17}$$

and T then is

$$T = 2\pi\sqrt{\frac{m}{k}}. \tag{9.18}$$

9.3.3 Further Analysis of Velocity and Acceleration of Mass-Spring System in SHM

A further analysis of the SHM can be done to address the positions, R, O, and L of Figure 9.4 or D, O, and H of Figure 9.5. The values of velocity, acceleration, kinetic energy, potential energy, and total energy of the mass as it passes these points during its cyclic motion are analyzed in the following discussion.

a. Point R: This point marks the farthest position to the right and corresponds to the maximum displacement. Then, the maximum force on m by the spring is

$$F_{net} = -kx_{max} = -kA. \tag{9.19}$$

From Newton's law,

$$F_{net} = ma.$$

Thus,

$$ma = -kA,$$

making the maximum acceleration of m,

$$a_{max} = -\left(\frac{k}{m}\right)A. \tag{9.20}$$

Using Equation 9.16, the maximum acceleration becomes

$$a = -\omega^2 A. \tag{9.21}$$

However, the velocity from Equation 9.9 is

$$v = \pm\omega\sqrt{(A^2 - x^2)} = \pm\omega\sqrt{(A^2 - A^2)} = 0.$$

Thus, the kinetic energy of the system at x = A is

$$K = \frac{1}{2}mv^2 = 0$$

and its potential energy at x = A is

$$U = \frac{1}{2}kx^2 = \frac{1}{2}kA^2. \tag{9.22}$$

This is the maximum value m can have during its oscillation.
The total energy of the system is

$$E = K + U = 0 + \frac{1}{2}kA^2 = \frac{1}{2}kA^2. \tag{9.23}$$

b. Point O: This is the equilibrium position for the mass and corresponds to its minimum displacement (x = 0). The force on m is

$$F_{net} = -kx_{min} = -k(0) = 0.$$

As from Newton's law,

$$F_{net} = ma,$$

then

$$ma = 0.$$

Hence, the acceleration a = 0, representing the minimum acceleration experienced by m.
However, the velocity from Equation 9.9 is $v = \pm\omega\sqrt{(A^2 - x^2)}$.
That is,

$$v_{max} = \pm\omega\sqrt{(A^2 - 0^2)} = \pm\omega A, \tag{9.24}$$

representing the maximum value for v. The negative sign for v is taken when m passes through O moving to the left, and the positive sign is taken when it passes through O moving to the right.

From Equation 9.24, as $v_{max} = \pm\omega A$ at $x = A$, then the kinetic energy of the system at $x = 0$ is maximum,

$$K_{max} = \frac{1}{2}m\omega^2 A^2, \tag{9.25}$$

while its potential energy at $x = 0$ would be minimum and is given by

$$U = \frac{1}{2}kx^2 = \frac{1}{2}k(0)^2 = 0.$$

The total energy of the system is

$$E = K + U = \frac{1}{2}m\omega^2 A^2 + 0.$$

That is,

$$E = \frac{1}{2}m\omega^2 A^2. \tag{9.26}$$

c. Point L: This point marks the farthest to the left among the mass positions corresponding to its maximum displacement upward.

The force on m at point L would be maximum as in case (a). The only difference would be in the sign of the force being positive, simply because it is directed along the positive x axis. The acceleration also will be of the same value as in case (a), but of a positive sign. However, all other quantities would have the same values calculated in case (a).

EXAMPLE 9.4

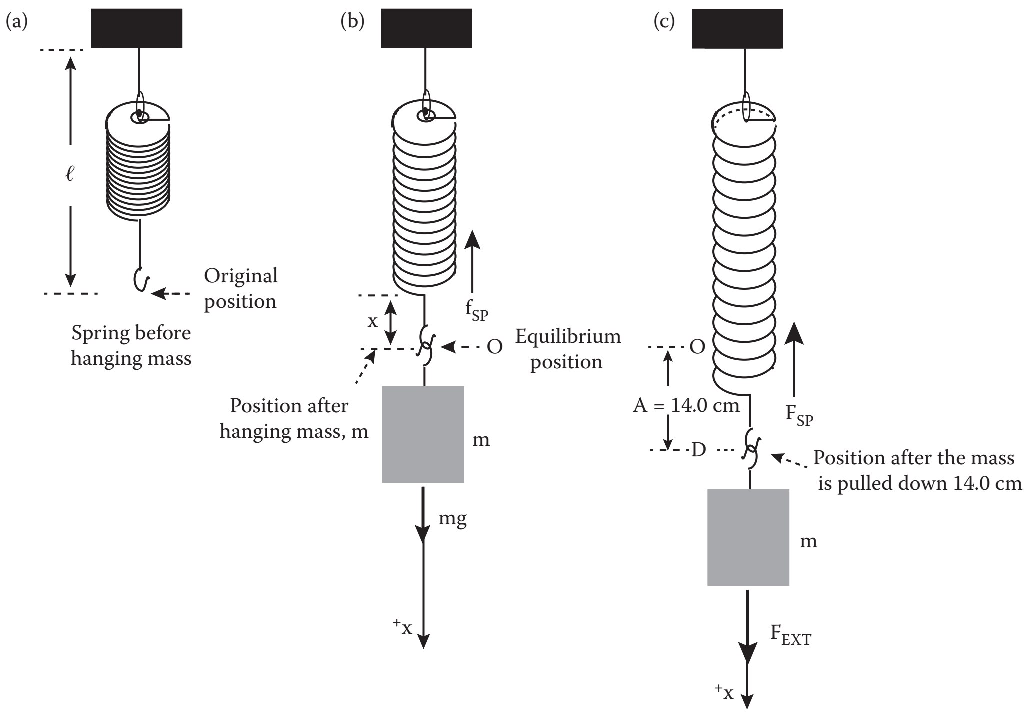
Consider the spring described in Example 9.3 ($k = 45.0$ N/m) that is attached to the ceiling (the following figure, part (a)) with the 0.64-kg block attached to its free end (part (b)). The block is pulled down to point D 14.0 cm from its equilibrium position before it was released (part (c)). Determine

a. The potential energy stored in the spring just before it is released
b. The kinetic energy of the block as it passes through its equilibrium position O
c. The system's total mechanical energy.
d. The velocity of the block at positions of the following displacements: (1) $x = 14.0$ cm, (2) $x = 11.0$ cm, block moving up, and (3) $x = 0.00$ cm.

Solution

a. From Equation 9.4,

$$U = \frac{1}{2}kx^2.$$



Thus,

$$U = \frac{1}{2}(45.0 \text{ N/m})(0.140 \text{ m})^2 = 0.44 \text{ J}.$$

As the 14-cm stretch is actually the maximum, which represents the amplitude of the oscillation, then the above value for U is the maximum value of the system's potential energy.

b. The kinetic energy at O, $(K)_o$, is maximum and is equal to the system's maximum potential energy obtained in part (a). Thus,

$$(K)_o = 0.44 \text{ J}.$$

c. The block's total mechanical energy E is conserved and is equal to its maximum kinetic energy. Thus,

$$E = U_{max} = K_{max} = 0.44 \text{ J}.$$

d. The velocity of the block at any position is given by Equation 9.9

$$v = \pm\omega\sqrt{(A^2 - x^2)},$$

where from Equation 9.8

$$\omega = \pm\sqrt{\frac{k}{m}} = \sqrt{\frac{45.0 \text{ N/m}}{0.64 \text{ kg}}} = 8.4 \text{ rad/s}.$$

As the amplitude of the oscillation A is 14.0 cm, then

1. For x = 14.0 cm, $v = \pm\omega\sqrt{(A^2 - x^2)} = \pm 8.4\ \text{rad/s}\sqrt{[(0.140\ \text{cm})^2 - (0.140\ \text{cm})^2]}$ = 0.00 m/s
2. For x = 11.0 cm, $v = \pm\omega\sqrt{(A^2 - x^2)} = \pm 8.4\ \text{rad/s}\sqrt{[(0.140\ \text{cm})^2 - (0.110\ \text{cm})^2]}$ = −0.73 m/s
3. For x = 0.00 cm, $v = \pm\omega\sqrt{(A^2 - x^2)} = \pm 8.4\ \text{rad/s}\sqrt{[(0.140\ \text{cm})^2 - (0.00\ \text{cm})^2]}$ = 1.2 m/s

Analysis

1. The above values show that the velocity of the block is greatest at O as it should be, and its velocity at its lowest point L is zero, also as expected.
2. The numerical value of v for the block at x = 11 cm on its way up is less than that at O and larger than its value at D, as expected.

Question

Determine (a) the period and (b) the frequency of the oscillation of the block in Example 9.4.

9.4 SIMPLE PENDULUM IN SHM

A simple pendulum is a bob of mass m attached to an inextensible string that is suspended from a solid support, C (Figure 9.6). At the mass equilibrium position, N, the string hangs vertically down. Once the bob is pulled through a small angle to a point O on the right side of N and then released, it moves under gravity back toward N, gaining momentum upon approaching equilibrium. Thus, the bob goes on to the other side of the vertical line, moving against gravity with a fading velocity. As the bob's velocity becomes zero at O′, it then gets pulled back by gravity toward equilibrium, approaching it with gradually increasing velocity. It then passes the equilibrium position and keeps moving to the right against gravity, diminishing in velocity until it becomes zero at O again. The bob then moves back under gravity, oscillating back and forth in identical cycles.

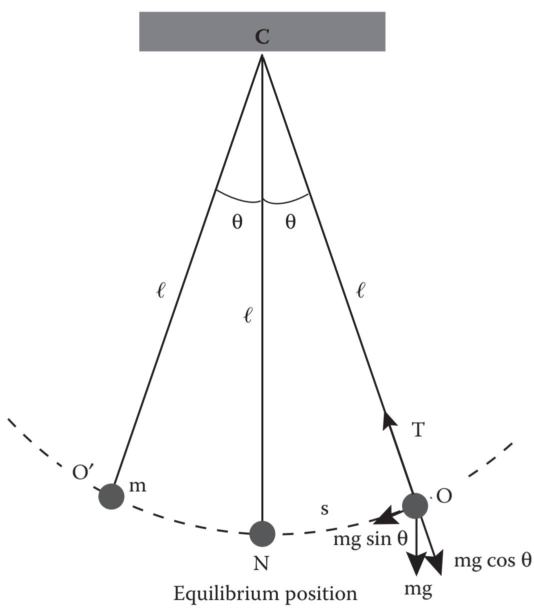


FIGURE 9.6 A simple pendulum that consists of a string of length ℓ attached to a rigid support on one end and to a mass m on its other end.

The analysis of the dynamics of the simple pendulum resides in defining the force that acts on the bob to restore it to its equilibrium position. This force is the component of the bob's weight along the tangent to the arc s. Denoting this component by F_t, then

$$F_t = -mg \sin \theta. \tag{9.27}$$

The minus sign is because the tangential force F_t is pointing toward the equilibrium position, N, while the angle θ is measured from the equilibrium line CN, increasing as the bob moves away from N. If the angle θ is small (<15°) such that $\sin \theta \sim \theta$ in radians, Equation 9.27 becomes

$$F_t = -mg\theta. \tag{9.28}$$

If θ is in radian, it subtends an arc s given by

$$s = \ell\theta.$$

Thus, Equation 9.28 becomes

$$F_t = -mg\frac{s}{\ell}, \tag{9.29}$$

which shows that the force acting on m along the arc is proportional and opposite to the displacement s. From Newton's second law, the tangential force

$$F_t = ma_t, \tag{9.30}$$

where a_t is the acceleration along the tangent to the path; that is, along the arc s. For rotational motion, the tangential acceleration, established in Chapter 8 (see Equation 8.9, Section 8.2), is

$$a_t = \ell\alpha, \tag{9.31}$$

where α is the angular acceleration of the pendulum's bob about C. Substituting Equation 9.31 into Equation 9.30 and using Equation 9.28 gives

$$m\ell\alpha = -mg\theta.$$

The above relation can now be written as

$$\alpha = -\frac{g}{\ell}\theta. \tag{9.32}$$

For the factor g/ℓ notices that it has the units of $(m/s^2)/m = s^{-2}$, then g/ℓ can appropriately be set equal to ω^2, where

$$\omega = 2\pi f \tag{9.33}$$

would then be the angular frequency of the mass m in its SHM along the arc s. Thus, Equation 9.32 becomes

$$\alpha = -\omega^2\theta, \tag{9.34}$$

where

$$\omega = \sqrt{\frac{g}{\ell}} \tag{9.35}$$

and

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{g}{\ell}}. \tag{9.36}$$

The period of the oscillation, defined as the inverse of the frequency, is

$$T = 2\pi\sqrt{\frac{\ell}{g}}. \tag{9.37}$$

Equation 9.37 shows that the period of the oscillation of a simple pendulum depends on the length of the pendulum, ℓ. Therefore, such a pendulum could be a useful measure of time. The other important observation in this equation is that it embraces the gravitational acceleration g. This makes the simple pendulum of a dual benefit. Measuring the period T can help determine the gravitational acceleration, g, which is a measure of the strength of the earth's gravitational field. Thus, Equation 9.37 has a more subtle implication that a simple pendulum may be used to measure the strength of the gravitational field of any planet on which such an experiment could be carried out. For example, the period of a simple pendulum on the Moon relative to the pendulum's period on the Earth would be of the value

$$\frac{T_M}{T_E} = \sqrt{\frac{g_E}{g_M}}. \tag{9.38}$$

EXAMPLE 9.5

Determine (a) the period, (b) the frequency, and (c) the angular frequency of a simple pendulum of length $\ell = 0.40$ m.

SOLUTION

a. From Equation 9.37, $T = 2\pi\sqrt{\ell/g}$, which upon substituting for $\ell = 0.400$ m and $g = 9.80\ \text{m/s}^2$ gives

$$T = 2\pi\sqrt{\frac{0.400\ \text{m}}{9.80\ \text{m/s}^2}} = 1.27\ \text{s}.$$

b. The frequency $f = 1/T$. Thus,

$$f = 1/1.27\ \text{s} = 0.787\ \text{s}^{-1}.$$

c. From Equation 9.11, the angular frequency is $\omega = 2\pi f$. Thus,

$$\omega = 2\pi(\text{rad})(0.787\ \text{s}^{-1}) = 4.95\ \text{rad/s}.$$

An analysis parallel to that described for a mass-spring system in Section 9.3.2 can be presented for a simple pendulum. The main difference is that the displacement of the pendulum is angular, while in the mass-spring system, it is linear. Thus, for a simple pendulum, the angular displacement from equilibrium is θ and its amplitude is the largest value, θ_o, through which the pendulum's bob is pulled to the side before it is released. The periodic function expressing θ as a function of t can take the form

$$\theta = \theta_o \cos(\omega_o t), \tag{9.39}$$

where ω_o is the natural angular frequency of the pendulum and

$$\omega_o = 2\pi f_o, \tag{9.40}$$

where f_o is the natural linear frequency of the pendulum's SHM. As noted earlier, this is equal to the inverse of the period of oscillation T_o. That is,

$$f_o = \frac{1}{T_o}. \tag{9.41}$$

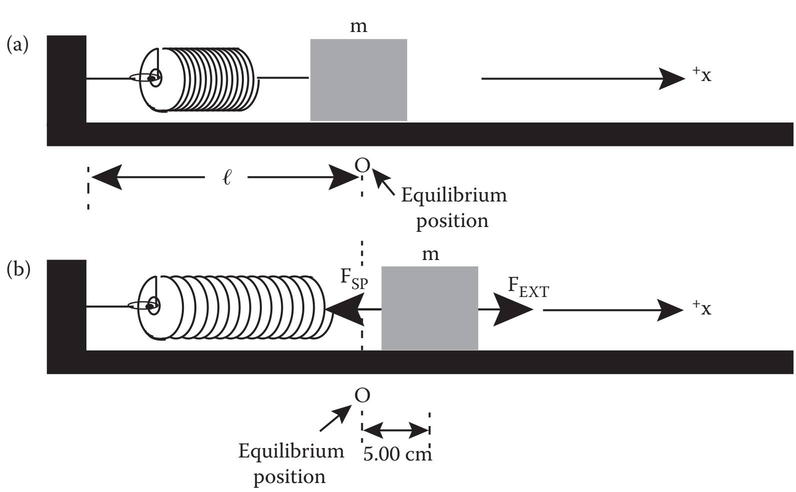
The form in Equation 9.39 is based on starting the oscillation when $\theta = \theta_o$.

PROBLEMS

9.1 A spring of stiffness constant k = 18.0 N/m and natural length ℓ is suspended vertically. If a ball of 0.22 kg mass is attached to the spring's free end, find how far down the mass stretches the spring?

9.2 In the previous problem, if the ball is pulled further, 6.0 cm below the equilibrium position and then let go, find
   a. The frequency of oscillation of this system
   b. The period of oscillation
   c. The velocity of the ball as it goes through the equilibrium point

9.3 A spring of stiffness constant k = 24.0 N/m and natural length ℓ = 28.0 cm, attached from one end to a side vertical solid support (the figure below), has a block of mass m = 0.42 kg attached to its other end. The spring is laid out on a horizontal frictionless table.
   a. How much work is needed to stretch the spring 5.0 cm further?
   b. Determine the frequency of the mass spring system as the mass is released after the 5.0-cm stretch.



9.4 A spring suspended vertically is found to stretch 24.0 cm when a 0.45-kg ball is attached to its free end. Determine
   a. The spring's stiffness constant k
   b. The work required to stretch the spring 3.30 cm from its equilibrium position
   c. The period of the oscillation the mass-spring system executes once it is released after the 3.3-cm extension

9.5 A 0.12-kg ball is suspended from a spring of stiffness constant k = 14.6 N/m. The ball is pulled down 4.40 cm from its equilibrium position. Find
   a. The angular frequency of oscillation
   b. The velocity of the ball at x = 4.40 cm and at the equilibrium position
   c. The velocity at x = 2.2 cm

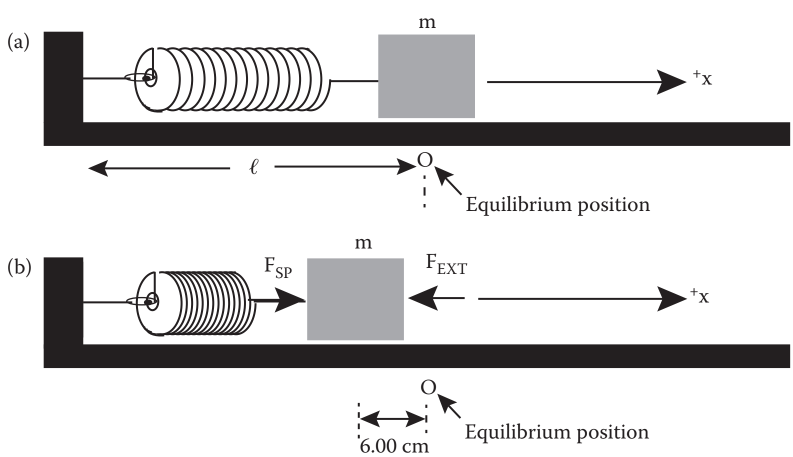
9.6 For the system in the previous problem, find
   a. The ball's maximum kinetic energy
   b. The spring's maximum potential energy
   c. The system's total mechanical energy

9.7 A block of mass m = 0.200 kg, attached to a vertically mounted spring, undergoes a harmonic oscillation of the form $x = A \cos(4\pi t)$; A = 4.00 cm. Determine the folllowing:
   a. The initial position of the block, that is, at t = 0
   b. The frequency f and period T of the motion
   c. The velocity of the block at t = T/2
   d. The block's maximum acceleration

9.8 In Problem 9.7, find the following:
   a. The acceleration of the block at t = 0.0 s, t = (T/4) s, and t = (3T/4) s
   b. The force acting on the block at t = 0.0 s, t = (T/4) s, and t = (3T/4) s
   c. The velocity of the block at t = 0.0 s, t = (T/4) s, and t = (3T/4) s

9.9 Use the information given in the previous problem and the answers to parts (a)–(c) to calculate (a) the kinetic, (b) potential, and (c) total mechanical energy of the mass-spring system at t = 0.0 s, t = (T/4) s, and t = (3T/4) s.

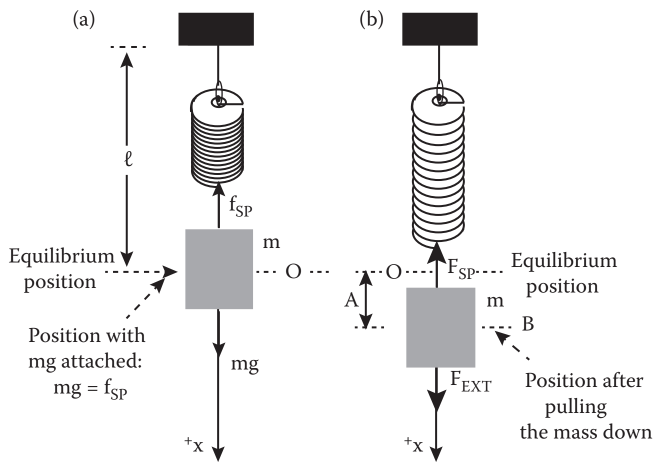
9.10 A spring, at rest on a frictionless table top, is attached to a rigid wall at one end and to 1.20-kg block at its other end (the figure below, part (a)). The spring is of stiffness constant k = 84.0 N/m. It is compressed 6.00 cm from its equilibrium position (part (b)) and then released to execute an SHM.



a. Determine the wave equation of the executed harmonic oscillation.
b. Determine the block's speed and acceleration when it passes through the equilibrium position.
c. Use the wave form established in part (a) to determine the position of the block at $t = 0.0$ s, $t = (T/8)$ s, and $t = (T/4)$ s.

9.11 A block of mass $m = 0.180$ kg, suspended from a spring of stiffness constant 18.0 N/m (the figure below, part (a)), is undergoing an SHM (part (b)) with total mechanical energy of 0.060 J. Determine

a. The amplitude A of the oscillation
b. The maximum velocity the mass would have during its motion
c. The linear and the angular frequency of the oscillation



9.12 In a laboratory experiment, a simple pendulum is observed to swing executing 20 oscillations in exactly 28.0 s. Determine

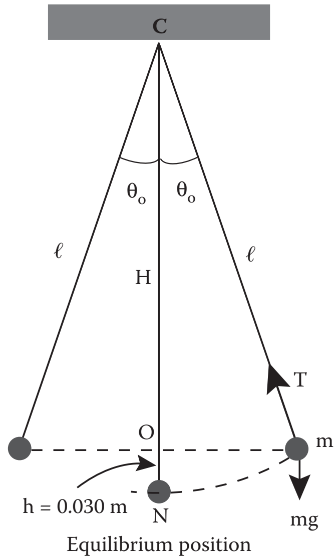
a. The period of oscillation
b. The frequency of oscillation
c. The length of the pendulum

9.13 If a simple pendulum identical to that described in the previous problem is taken to the Moon where the acceleration of gravity is 1.63 m/s^2, determine

a. The period of oscillation
b. The frequency of oscillation

9.14 A simple pendulum of a length $\ell = 90.0$ cm has a bob of mass $m = 0.150$ kg. The bob is pulled aside (the following figure) raising it to a height of 3.00 cm above its equilibrium position before it is released. Determine

a. The potential energy of the pendulum when the bob is at its highest position on either side of its equilibrium position
b. The kinetic energy of the pendulum when the bob is at its lowest position
c. The velocity of the bob at its lowest position

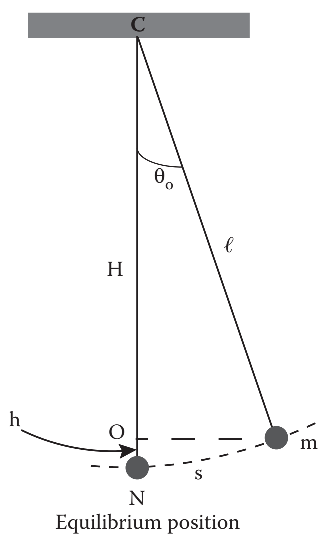


9.15 For the simple pendulum in the previous problem,

a. Determine the angular velocity of the harmonic oscillation executed by the pendulum
b. Through an intuitive parallel to the simple hamonic motion executed by an ideal spring, write down the wave equation for the oscillation of the pendulum
c. The phase constant of the motion

9.16 A simple pendulum of length $\ell = 60.0$ cm has a bob of mass $m = 0.120$ kg. The bob is pulled aside where it was raised to a height h above its equilibrium position before it is released (the figure below). If the timing of the oscillation is started when the bob is at its lowest position, moving with a velocity of 1.00 m/s, determine

a. The height above the equilibrium to which the bob is initially raised
b. The amplitude of the wave
c. The phase constant of the motion
d. The wave equation of the oscillation



10 Thermal Physics
Temperature, Heat, and Thermal Expansion

This chapter presents a new area of physics titled as thermal physics. Although heat as a concept related to hot objects is an old concept, the essence of thermal physics started to become well defined only at the turn of the nineteenth century when new notions and physical quantities relevant to heat were developed. Among these were internal energy, entropy, and enthalpy. Heat is now commonly perceived as a form of random energy that is transferred from hot objects to cold objects when these objects are in thermal contact. This perception came as a result of pioneering experiments by Joule who demonstrated in a quantitative manner the equivalence between heat and energy. The exchange of heat from a hot object to a cold one ceases when the two objects get to a state of equilibrium. Work on this field leads to a formulation of two laws of thermal physics known as the first and second laws of thermodynamics. For the sake of brevity, this chapter will be limited to the discussion of heat with a special focus on expansion of solids.

10.1 HEAT IS A FORM OF ENERGY, SPECIFIC HEAT, AND HEAT CAPACITY

James Joule (1818–1889) was the first to brilliantly demonstrate the equivalence of mechanical work and heat. He arrived at a value for the conversion factor between work in units of Joules (J) done on a system and the amount of heat in units of calories (cal) delivered to the system. This factor is expressed by

$$1.000 \text{ cal} = 4.186 \text{ J}. \tag{10.1}$$

The calorie is defined as the amount of heat necessary to raise the temperature of 1 g of water by 1.0°C from 14.5°C to 15.5°C. A relaxed definition of the calorie as the amount of heat energy necessary to raise the temperature of 1.0 g of water by 1.0°C is commonly used. The variation in temperature on the definition of the cal for water in the temperature range from 15°C to 90°C is limited to about 0.4%, and hence this change can be ignored. The quantity 1.00 cal/g °C (=1.00 kcal/kg °C) or 4186 J/kg °C is known as the specific heat of water, and this is denoted by c. The specific heat of a material is then the amount of heat necessary to raise the temperature of 1 g of that material by 1.0°C. The specific heat differs from one substance to another, and hence it is a unique thermal quantity for any material. It is also worth noting that specific heat is subject to the conditions of the substance as to whether it is kept under constant volume or constant pressure, an issue that is critical for gases. For liquids and solids, such a change is not of a significant concern, and specific heat under constant pressure is more practical to employ and is commonly used. Table 10.1 lists the specific heat for some materials under constant pressure and at 25°C. For an object of mass m, the amount of heat, C, required to raise its temperature by 1°C is

$$C = mc, \tag{10.2}$$

where c is the specific heat of the substance from which the object is made. The quantity C is called the heat capacity of the object, and as can be seen is a property of the object. Therefore, if the

TABLE 10.1
Specific Heat of Selective Substances, Solids and Liquids, at Atmospheric Pressure (1 atm)

Substance	Specific Heat (cal/g °C or kcal/kg °C)	Specific Heat (kJ/kg °C)
Aluminum	0.215	0.900
Copper	0.092	0.385
Gold	0.031	0.130
Iron	0.108	0.452
Lead	0.031	0.130
Silver	0.057	0.239
Glass (fused quartz)	0.18	0.753
Water[a]	0.998	4.18
Ice (−4.5°C)	0.498	2.08
Titanium	0.125	0.523
Platinum	0.032	0.134
Nickel	0.106	0.444
Mercury	0.033	0.138

Source: *CRC Handbook of Chemistry and Physics*, 70th Edition, 1989–1990, pages D-187, D-174, D-173, F-64. With permission.

[a] The well-known value of 1.00 cal/g °C for the specific heat of water is at 14°C.

temperature of an object is raised from an initial temperature T_i to a final temperature T_f, the heat energy Q gained by the object is

$$Q = mc\,(T_f - T_i) \tag{10.3}$$

or

$$\Delta Q = mc\Delta T. \tag{10.4}$$

Comment:

It is important to know that the cal unit for heat is different from the food calorie unit denoted by Cal that is commonly observed on labels of food products. One food calorie, Cal, is equal to 1000 heat calories.

EXAMPLE 10.1

Calculate the amount of heat needed to raise the temperature of a 1.20-kg block of iron from 10°C to 25°C.

Solution

From Table 10.1, the specific heat of iron is c = 0.108 cal/g °C and the temperature difference is ΔT = 15°C. Using Equation 10.4, the required amount of heat is

$$\Delta Q = mc\Delta T.$$

Thus, ΔQ = (1.20 kg)(0.108 kcal/kg °C) (15°C).
This gives ΔQ = 1.9 kcal.

10.2 HEAT AND INTERNAL ENERGY

Heat is energy that may be transferred from one system to another. This transfer occurs when there is a difference in temperature between the two systems. A typical symbol for the amount of heat energy exchanged is Q. When heat is gained by a system, its internal energy is changed. This kind of energy relates to the random motion of the system's atoms and molecules, which in turn is directly related to its temperature. The analogy between the heat transferred to a system and the change in its internal energy is similar to the connection between the work done on or by the system and the change in the system's kinetic energy.

10.3 ZEROTH LAW OF THERMODYNAMICS

The zeroth law of thermodynamics states that if two objects, A and B, are independently in thermal equilibrium with a third object, C, then the two objects, A and B, are in equilibrium with each other. In other words, thermal equilibrium between two objects implies that the two objects are at the same temperature. In this context, one can envision the function of a thermometer. A thermometer acts as a device made to measure an object's temperature once the two, the object and the thermometer, are brought to a thermal equilibrium with each other. At that instant, the thermometer indicates the temperature of the object. The zeroth law of thermodynamics requires two concepts that are of relevance to the exchange of heat. These are thermal contact and thermal equilibrium that are defined as follows:

1. *Thermal contact*: Two objects initially at different temperatures, placed in contact with each other, exchange heat energy.
2. *Thermal equilibrium*: Thermal equilibrium is a state of two objects in contact with each other but with no heat energy exchanged between them.

10.4 THERMOMETERS AND TEMPERATURE SCALES

As stated above, the concept of thermal equilibrium is utilized in having thermometers measure the temperature of objects. Without going into the details of how thermometers are calibrated and what substances are used in them, there are three scales (Figure 10.1) that are described as follows.

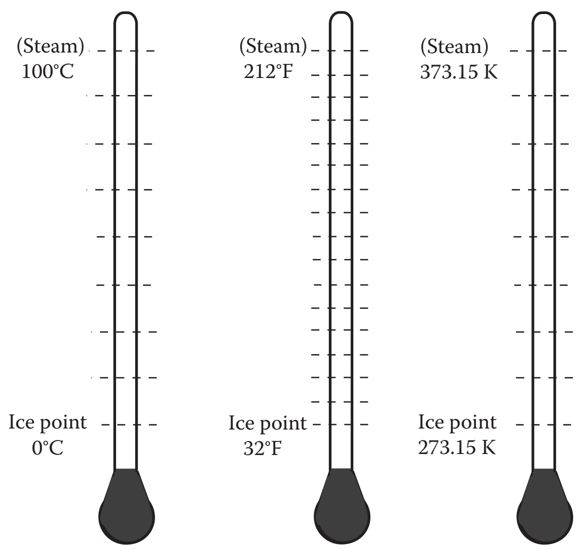


FIGURE 10.1 Three thermometers scaled in Celsius, Fahrenheit, and Kelvin.

10.4.1 Celsius Temperature Scale (°C)

This scale takes the temperature of ice–water mixture as the 0°C, while the temperature of the water–steam mixture is defined as 100°C.

10.4.2 Fahrenheit Scale (°F)

This scale is used in the United States and England. The size of each degree on this scale is different from the Celsius scale. This scale defines the temperature of ice–water mixture to be 32°F, while taking the temperature of the water–steam mixture to be 212°F. The range between these two temperature points is 180°F. A comparison of the Fahrenheit range to the 100°C range on the Celsius scale gives the value (180/100). Thus, the ratio between the sizes of one degree on the two temperature scales is $(\Delta T_F/\Delta T_C) = (9/5)$. This offers a conversion factor between a reading on the Fahrenheit scale to the corresponding value on the Celsius scale, which when we take care of accounting for the 0°C = 32°F becomes

$$T_F = \frac{9}{5}T_C + 32°F \tag{10.5}$$

or, equivalently,

$$T_C = \frac{5}{9}(T_F - 32°F). \tag{10.6}$$

10.4.3 Kelvin Temperature Scale (K)

Thermometers marked with a Kelvin scale are also called constant-volume gas thermometers. This scale is related to the Celsius scale according to the definition

$$T_C = T_K - 273.15. \tag{10.7}$$

In the Celsius and Kelvin scales, the size of each degree is the same. Only the reference points differ. As Equation 10.7 shows, 0°C on the Celsius scale corresponds to 273.15 K on the Kelvin scale.

EXAMPLE 10.2

Calculate the quantity of heat in cal, and in Joules, required to raise the temperature of each of the following substances from 30.0°C to 82.0°C.

a. 22.0 g water
b. 22.0 g gold
c. 22.0 g aluminum

Solution

From Equation 10.4 $\Delta Q = mc\Delta T$.

a. Substituting for the given quantities for water gives:
   In calories, $\Delta Q = (22.0\ \text{g})(0.998\ \text{cal/g °C})(52.0°\text{C}) = 1140\ \text{cal}$.
   In Joules, $\Delta Q = (1140\ \text{cal})(4.186\ \text{J/cal}) = 4770\ \text{J}$.
b. Substituting for the given quantities for gold gives:
   In calories, $\Delta Q = (22.0\ \text{g})(0.031\ \text{cal/g °C})(52.0°\text{C}) = 36\ \text{cal}$.
   In Joules, $\Delta Q = (36\ \text{cal})(4.186\ \text{J/cal}) = 150\ \text{J}$.

c. Substituting for the given given quantities for aluminum gives:
   In calories, $\Delta Q = (22.0\ g)(0.215\ cal/g\ °C)(52.0°C) = 250\ cal$.
   In Joules, $\Delta Q = (250\ cal)(4.186\ J/cal) = 1100\ J$.

Analysis

It should be noted that answers in parts (b) and (c) were rounded off to two significant figures. It is noticed that the difference in the amount of heat needed to raise equal masses of three different substances, initially at the same temperature, to the same final temperature is widely different. Among these three, the amount of heat is the lowest for gold and the highest for water.

EXAMPLE 10.3

The temperature on a typical hot summer day in Pennsylvania is 84°F. Determine the corresponding temperature on a Celsius scale.

Solution

From Equation 10.6

$$T_C = \frac{5}{9}(T_F - 32°F).$$

Thus,

$$T_C = \left(\frac{5}{9}°C/°F\right)(84°F - 32°F) = 29°C.$$

EXAMPLE 10.4

The boiling point of liquid nitrogen is at 77.0 K. Determine the corresponding value in (a) Celsius and (b) Fahrenheit.

Solution

a. From Equation 10.7

$$T_C = T_K - 273.15.$$

Thus,

$$T_C = 77.0\ K - 273.15\ K = -196°C.$$

b. From Equation 10.5

$$T_F = \left(\frac{9}{5}°F/°C\right)(-196°C) + 32.0°F = -321°F.$$

10.5 THERMAL EXPANSION

In general, heating metals causes them to expand. For a regularly shaped object that has length, width, and height, the increase occurs along all these dimensions. Therefore, the object experiences a change in its length, area, and volume.

10.5.1 Linear Expansion

In the case of a thin wire-like solid whose initial temperature T_o has changed to a final temperature T, a change in the wire's length occurs according to the following empirical relation:

$$\Delta L_o = \alpha L_o \, \Delta T, \tag{10.8}$$

where L_o is the length of the wire at temperature T_o, α is the coefficient of linear expansion of the wire, and $\Delta T = T - T_o$. Table 10.2 lists the linear coefficient of expansion, α, for some solids. As can be analyzed from the dimensions of the terms in Equation 10.8, the unit for α can be in °C^{-1} or °F^{-1}. In addition, the values of α for most solids are of the order of 10^{-6} °C^{-1}.

EXAMPLE 10.5

A thin copper rod 16.00 cm long, initially at room temperature 23.0°C, is heated to a temperature of 55.0°C. Calculate the new length of the rod.

Solution

From Equation 10.8

$$\Delta L_o = \alpha L_o \Delta T.$$

Substituting for the given quantities gives:

$$\Delta L_o = (1.7 \times 10^{-5} \text{ °C}^{-1})(16.00 \text{ cm})(32.0\text{°C})$$

$$= 8.7 \times 10^{-3} \text{ cm} = 0.087 \text{ mm}.$$

The rod's new length becomes $L = L_o + \Delta L_o = 16.00 + 0.0087 = 16.01$ cm.

Analysis

Notice that to keep the proper number of significant figures, the second decimal place was rounded to 1. The change in length is in fact <0.01 cm.

TABLE 10.2
A List of the Linear Coefficient of Expansion for Some Selective Solids

Substance	Coefficient of Linear Expansion ($\times 10^{-6}$) (°C^{-1})
Aluminum	25
Gold	14.2
Copper	16.6
Lead	29
Silver	19
Iron	12
Tin	20
Nickel	13
Cobalt	12
Tungsten	4.5

Source: *CRC Handbook of Chemistry and Physics*, 70th Edition, 1989–1990, p. D-187. With permission.

10.5.2 Surface Expansion

If an object has two dimensions such as a plate, rectangular, circular, or otherwise, the area of the object increases by increasing its temperature according to the following relation:

$$\Delta A_o = \gamma A_o\, \Delta T, \tag{10.9}$$

where γ is the coefficient of area expansion and A_o is the original area of the object. The units for γ are as those for α in °C^{-1} or °F^{-1}. A square plate of a side L_o at temperature T expands after a change in its temperature by ΔT so that each side would become $L = (L_o + \Delta L_o)$, and hence the plate's area, A, after expansion becomes

$$A = L^2, \quad L = L_o + \Delta L_o = L_o + \alpha\, L_o\, \Delta T.$$

Thus,

$$A_o + \Delta A_o = (L_o + \Delta L_o)^2 = (L_o + \alpha\, L_o\, \Delta T)^2$$

or

$$A_o + \Delta A_o = (L_o^2 + 2\alpha\, L_o^2\, \Delta T + \alpha^2\, L_o^2\, \Delta T^2).$$

Since α is a very small quantity, its square is practically negligible. As $A_o = (L_o)^2$, the above equation becomes

$$A_o + \Delta A_o = L_o^2 + 2\alpha\, L_o^2\, \Delta T = A_o + 2\alpha A_o \Delta T.$$

Thus,

$$\Delta A_o = 2\alpha A_o \Delta T.$$

Comparing the above equation with Equation 10.9 suggests that

$$\gamma = 2\alpha. \tag{10.10}$$

EXAMPLE 10.6

A thin copper plate 16.00 cm on a side, initially at room temperature 23°C, is heated to a temperature of 55°C. Calculate the new area of the square.

Solution

From Equations 10.9 and 10.10

$$\Delta A_o = \gamma A_o \Delta T = 2\alpha A_o \Delta T.$$

After substituting for the given quantities, ΔA_o becomes

$$\Delta A_o = (2 \times 1.7 \times 10^{-5}\ °C^{-1})\ (16.00\ cm)^2 (32.0°C) = 0.28\ cm^2.$$

The new area becomes

$$A = (16.00)^2\ cm^2 + 0.28\ cm^2 = 256.0\ cm^2 + 0.28\ cm^2 = 256.3\ cm^2.$$

Analysis

Notice that to keep the proper number of significant figures, the first decimal place was rounded to 3. The change in area is in fact <0.3 cm^2.

10.5.3 Volume Expansion

The increase in the volume of an object as a result of raising its temperature by an amount ΔT is

$$\Delta V_o = \beta V_o \Delta T, \tag{10.11}$$

where

$$\beta = 3\alpha \tag{10.12}$$

is the coefficient of volume expansion and V_o is the original volume of the object. The units of β are $°C^{-1}$ or $°F^{-1}$.

Comment:

1. Following the discussion through which we established that the surface coefficient of expansion γ was two times the linear coefficient of expansion α, one could lay out a similar argument and derive Equation 10.12.
2. Equations 10.9 through 10.12 remain valid for any two- or three-dimensional object that has a portion cut from it. The cut that leaves the object with a hole does not impact the way the object is treated.

Each of the coefficients, α, γ, and β, is unique for each material. That is, their values for a certain material are characteristics of that material.

EXAMPLE 10.7

A solid copper cube of 16.00 cm on a side and initially at room temperature 23°C is heated to 55°C. Calculate the new volume of the cube.

Solution

From Equation 10.11

$$\Delta V_o = \beta V_o \Delta T = 3\alpha V_o \, \Delta T.$$

After substituting for the given quantities,

$$\Delta V_o = (5.1 \times 10^{-5}\ °C^{-1})(16.00\ \text{cm})^3(32.0°C) = 6.7\ \text{cm}^3.$$

The new volume becomes

$$V = (16.00)^3\ \text{cm}^3 + 6.7\ \text{cm}^3 = 4096\ \text{cm}^3 + 6.7\ \text{cm}^3 = 4102.7\ \text{cm}^3.$$

Analysis

As the original volume V_o is 4096 cm^3, the change in the volume turns out to be 6.7 cm^3, which is <0.16% of the original volume.

10.6 CALORIMETRY

Calorimetry is a quantitative analytical procedure that is used to determine the specific heat of a solid. This is done by heating the solid to a known temperature and immersing it directly into a known amount of water at an initially known temperature, usually room temperature. Once thermal equilibrium between the water and the solid is achieved, the final temperature of the combination will be of a value somewhere between that of the solid before immersion and the initial temperature of water. Assuming that both the confined water and the solid are isolated from all surroundings, measurement of the final temperature implies that the heat lost by the solid should be equal to the heat gained by the water. This technique is called calorimetry, and the vessels that are used for this purpose are called calorimeters. Figure 10.2 is a sketch of a typical simple calorimeter. The inner part of the calorimeter has to be thermally isolated from all surroundings. Such a condition is secured in a vessel that consists of a double-wall container with a tight lid on it that allows for inserting a thermometer to measure the temperature of the enclosure. The two walls of the calorimeter are isolated with air in-between or by a thermally synthetic insulating substance filling the space between the two walls.

The concept of calorimetry is based on the principle of conservation of energy, and since the system (water, solid, and calorimeter) is assumed completely insulated, exchange of heat energy between the two objects, one hot and the other cold, requires that the following condition holds:

Heat gained by the cold object = Heat lost by the hot object.

In symbols, the above relation becomes

$$\Delta Q_{gained} = -\Delta Q_{lost}. \tag{10.13}$$

The process goes as follows. Consider a piece of metal that has unknown specific heat, c_s (c_s stands for solid) that is heated to a relatively high predetermined temperature, T_i, and let its mass be m. If this piece is gently immersed in an amount of water of mass m_w, that is at a lower T_w, then

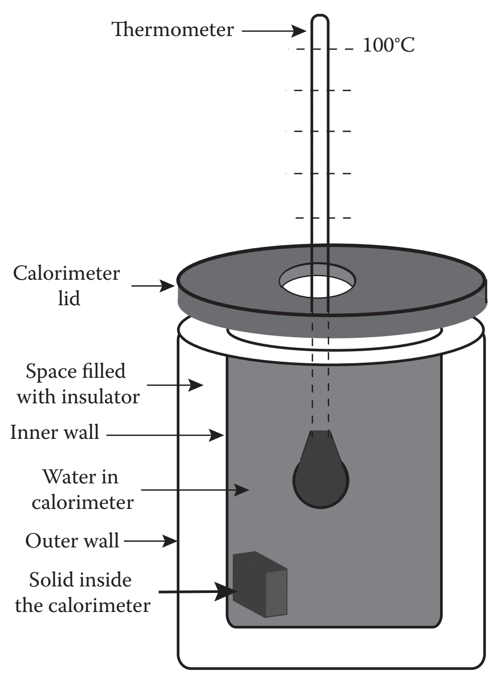


FIGURE 10.2 A calorimeter consisting of a double-wall container, inner and outer, isolated from the surroundings.

after waiting for thermal equilibrium to be reached, the final temperature of the whole enclosure would be of a value higher than that of water and lower than that of the metal piece.

Letting the final temperature of the whole be T_f, and labeling the specific heat of water and the calorimeter by c_w and c_c, respectively, then from Equation 10.13 the heat energy gained by the contained water and the calorimeter is

$$\Delta Q_{w\ gained} = m_w\, c_w \Delta T_w + m_c\, c_c \Delta T_c, \tag{10.14a}$$

where

$$\Delta T_c = \Delta T_w = T_f - T_w. \tag{10.14b}$$

Notice that the calorimeter and the water are initially at the same temperature, and both get to the same final temperature.

The heat energy lost by the metal piece is

$$\Delta Q_{\ lost} = m\, c_s \Delta T_s, \tag{10.15a}$$

where

$$\Delta T_s = T_f - T_i. \tag{10.15b}$$

Applying Equation 10.13 and using Equation 10.15a gives

$$[m_w\, c_w\, \Delta T_w + m_c\, c_c\, \Delta T_c] = -m\, c_s\, \Delta T_s. \tag{10.15c}$$

Since all quantities except c_s are known, the specific heat of the metal, c_s, can be determined.

Remark 10.1

It is important to note that ΔT_w, ΔT_c, and ΔT_s. Each of the quantities on each side of Equation 10.15c has been dealt with as the magnitude of heat energy gained by elements of the left side and lost by the element on the right side.

EXAMPLE 10.8

An empty cup of mass 0.400 kg and temperature of 23°C was filled with 0.100 kg of boiling water. The final temperature of the water filled cup after equibrium was 67°C. Determine the specific heat of the cup material.

Solution

In this problem, only the cup gained heat, and that was the amount lost by the boiling water. From Equation 10.13

$$\Delta Q_{gained} = -\Delta Q_{lost}.$$

Using Equation 10.15c, after dropping the calorimeter term, becomes

$$m_{cup} c_{cup} \Delta T_{cup} = -m_w c_w \Delta T_w$$

$$(0.400\ \text{kg})(c_{cup})(67 - 23°\text{C}) = -(0.100\ \text{kg})(1.00\ \text{kcal/kg °C})(67 - 100°\text{C}).$$

This gives

$$c_{cup} = 0.19\ \text{kcal/kg °C}.$$

Analysis

Contrasting the above value against those in Table 10.1 suggests that the cup is closest to be made of glass.

EXAMPLE 10.9

A block of iron of 0.210 kg whose specific heat is 0.108 kcal/kg °C mass was heated to 495°C and dropped in an aluminum vessel of 95.0 g mass partially filled with 0.400 kg water. Knowing that the initial temperature of the vessel with the water was 21.0°C, find the value of the final temperature of the vessel, water, and iron block.

Solution

This problem seems to be a typical calorimetry problem in which the calorimeter and water are gaining heat, while the iron piece is losing it. From Equation 10.15c

$$m_w c_w \Delta T_w + m_c c_c \Delta T_c = -m_s\, c_s \Delta T_S.$$

With temperature differences on each side dealt with as magnitudes, the above equation becomes

$$(0.400\ \text{kg})(0.998\ \text{kcal/kg °C})(T_f - 21.0°\text{C}) + (0.0950\ \text{kg})(0.215\ \text{kcal/kg °C})(T_f - 21.0°\text{C})$$

$$= -(0.210\ \text{kg})\ (0.108\ \text{kcal/kg °C})\ (T_f - 495°\text{C}).$$

The above equation reduces to

$$0.399T_f - 8.38 + 0.0204T_f - 0.428 = -0.0227T_f + 11.2.$$

With some algebra, the above equation becomes

$$0.442T_f - 8.81 = +11.2.$$

Thus,

$$T_f = 45.3°\text{C}.$$

Analysis

Recalling that the mass of the iron piece being only 210 g, its initial temperature being so high was able to raise the whole enclosure to 45.6°C is significant.

10.7 HEAT TRANSFER: TRANSFER BY CONDUCTION

It has been discussed that heat transfer between two bodies occurs when there is a temperature difference between them. There are three processes of transferring heat between hot and cold objects. These are heat transfer by conduction, convection, and radiation. Heat transfer by convection is most common in liquids and gases, while heat transfer by radiation occurs when a hot object radiates energy in the form of electromagnetic waves, which get absorbed by nearby objects and are converted into heat.

The transfer of heat by conduction is most relevant to solids that are in a physical contact. This can best be demonstrated for a solid piece in the form of a slab of two ends at two different temperatures (Figure 10.3). Heat transfers through the atoms of the hot end to the neighboring atoms in the direction of the cold end. The process continues until thermal equilibrium is attained. In case the two ends are kept at two different temperatures T_1 and T_2 where $T_2 > T_1$, heat energy continues to flow until a thermal steady state is achieved. At this stage, every part of the slab retains the temperature it has reached. It was found that, at this stage, the rate of heat flow per unit cross-sectional

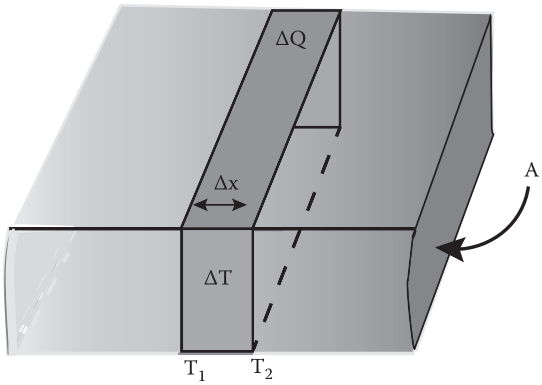


FIGURE 10.3 A schematic of a slab whose two ends are at two temperatures T_1 and T_2.

area of the slab is proportional to the temperature difference between two ends of the slab and is inversely proportional to the width of the slab. This can be expressed by the equation:

$$\frac{1}{A}\frac{\Delta Q}{\Delta t} = k\frac{(T_2 - T_1)}{\Delta x}, \tag{10.16}$$

where A is the cross-sectional area of the slab and k is called the thermal conductivity of the material from which the slab is made. Equation 10.16 can be rewritten as

$$\frac{\Delta Q}{\Delta t} = kA\frac{\Delta T}{\Delta x}, \tag{10.17}$$

where $\Delta T = T_2 - T_1$ is the temperature difference between the two cross sections at the two sides of the segment Δx. In the limit when the segment is taken to be infinitesimally small, the quantity $\Delta Q/\Delta t$ signifies the rate of heat transfer and is denoted by dQ/dt. This is also called the heat current, while the quantity $\Delta T/\Delta x$ becomes dT/dx and is known as the temperature gradient.

Substances differ in their thermal conductivity from being good heat conductors such as metals to low or poor heat conductors such as dielectrics. In addition, k is also temperature dependent. Table 10.3 lists values of thermal conductivity at 25°C for some metals.

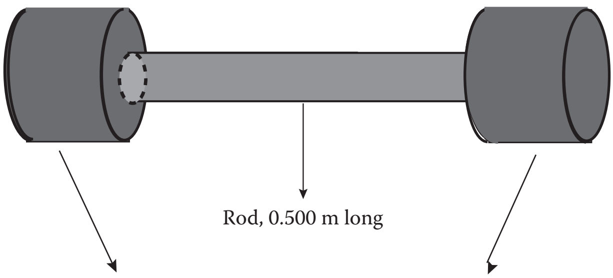
TABLE 10.3
Thermal Conductivity of Several Substances

Substance	Thermal Conductivity ($\times 10^2$) (W/m°C)
Aluminum	2.37
Copper	3.98
Iron	0.803
Gold	3.15
Silver	4.27
Lead	0.346
Nickel	0.899
Platinum	0.73
Titanium	0.2

Source: *CRC Handbook of Chemistry and Physics*, 70th Edition, 1989–1990, p. D-187. With permission.

EXAMPLE 10.10

Consider a copper bar 0.500 m long with 1.00 cm^2 (= $1.00 \times 10^{-4} m^2$) cross-sectional area. If the bar's two ends were placed in physical contact with two heat reservoirs that are at fixed temperatures of 18°C, 80°C (the figure below), what would be the value of the rate of heat transfer per second across the bar's cross section?



Solution

Substituting for the given quantities in Equation 10.17,

$$\frac{\Delta Q}{\Delta t} = kA\frac{\Delta T}{\Delta x}$$

gives

$$\frac{\Delta Q}{\Delta t} = \left(398\frac{W}{m°C}\right)(1.00 \times 10^{-4}\ m^2)\left(\frac{62.0°C}{0.50\ m}\right)$$

$$\frac{\Delta Q}{\Delta t} = 4.9\ W.$$

Analysis

As noticed, the heat transfer of ($\Delta Q/\Delta t$) = 4.9 W = 4.9×10^{-3} kW is a rather small quantity of power rate of heat transfer.

PROBLEMS

10.1 Calculate the amount of heat in (a) cal and (b) J needed to raise the temperature of 1.0 kg water at 23°C to its boiling point (100°C).

10.2 Convert the following temperatures to Celsius and Kelvins: (a) 32.0°F; (b) 72.0°F; (c) 212°F.

10.3 In the previous problem, use your answers of 0.0°C and 22.2°C to parts (a) and (b), respectively, to find the difference between them in
  a. Kelvin (K)
  b. Fahrenheit (°F)

10.4 Convert the following temperatures to Kelvins: (a) boiling point of liquid nitrogen −196.15°C; (b) boiling point of liquid helium −268.95°C.

10.5 At room temperature (23°C), a rod of an unknown material has a length L_o = 1.200 m and is found to increase by 1.00 mm when its temperature is raised to 95°C.
  a. Find the coefficient of linear expansion of the rod in $°C^{-1}$.
  b. Find the coefficient of linear expansion of the rod in $°F^{-1}$.
  c. Check Table 10.2 to give your best guess to what the rod's material would be.

10.6 A machinist, working on welding a copper ring of inner radius r = 6.000 cm onto a copper cylinder of radius R = 6.033 cm, has to have it barely slip over the cylinder first. If both originally are at room temperature, 23°C, find the temperature to which the ring should be raised so that it would match the radius of the copper cylinder.

10.7 The temperature of a spherical copper shell of 10.00 cm radius is raised from 23°C to 72°C. Determine the (a) radius, (b) surface area, and (c) volume of the shell at 72°C.

10.8 To a good approximation, the specific heat capacity of aluminum is twice that of steel. Consider two identical bars, one aluminum and another steel, each measures 1.0 m long at room temperature (23°C). Find the expansion of each bar in a hot day of 42°C.

10.9 A square plate of iron at room temperature (23°C) is 30.0 cm on a side. Calculate the percentage increase in the plate's area as the plate's temperature is increased to 285°C.

10.10 The mass density of silver at room temperature (23.0°C) is 1.05×10^4 kg/m^3,

a. Find the density of silver at 155.0°C. (*Hint*: The mass density in kg/m^3 is the mass of 1.000 m^3.)

b. Calculate the percentage change in a volume V_o of this silver.

10.11 Calculate the energy in (a) cal and (b) J that is contained in a 455-Cal piece of cheese cake. Find the height that a mountain climbed by a 65.0-kg man would result in having his gravtational potential energy increased by the amount calculated in part (b). (*Hint*: One food calorie, Cal, is equal to 1000 heat calories, cal.)

10.12 An aluminum tea pot whose mass when empty is 0.280 kg is filled with 0.600 kg water and both are at a room temperature of 23°C. Calculate the amount of heat to raise both to a temperature of 95°C.

10.13 In a laboratory experiment, an aluminum calorimeter of 0.111 kg mass is filled with 0.254 kg of water, all at room temperature (22°C). A slab 0.222 kg of copper, heated on a hot plate to a temperature of 255°C, is dropped gently in the calorimeter. Calculate the equilibrium temperature of the whole mix.

10.14 Consider a copper rod 2.50 m long with a circular cross section of 2.00 cm radius. The rod's ends are placed in a physical contact with two heat reservoirs that are at fixed temperatures of 22°C and 93°C. If the thermal conductivity of copper is 390 (W/m °C) and assuming it has reached thermal equilibrium, calculate

a. The temperature gradient

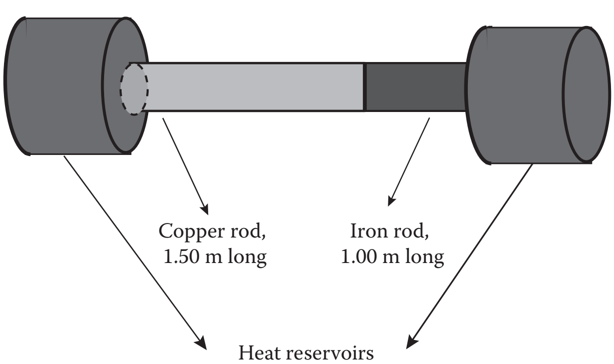
b. Rate of change of heat transfer across the cross sectional area of the rod

c. The temperature of the rod at its midpoint

10.15 Consider two rods, one is copper, 1.50 m long, and the other is iron, 0.600 m long (the figure below). Both rods are of identical circular cross section of 1.20 cm radius. The two rods are pressed firmly against each other. The other two ends are in physical contact with two heat reservoirs at fixed temperatures of 22°C and 93°C. Assuming the two-rod combination has reached a steady-state condition, calculate

a. The temperature at the area of contact

b. Rate of change of heat transfer across the cross-sectional area where the two rods are in contact



11 Waves and Wave Motion

A wave is a periodic disturbance in a medium. The medium can be a material as it is the case with all known waves other than light. A rope fastened at one end to a rigid support, and manually moved up and down from its other end, generates a pulse that could develop into a wave form travelling along the rope. A string held fixed at one end and an oscillator acting at the other end generates pulses that travel along the string. The waves getting reflected at the fixed end, and combining with the oscillator-generated ones, produce standing waves. In ideal circumstances, a wave consists of identically repeated pulses, each of which is called a cycle. The cycle is characterized by the following fundamental quantities: its duration that is called period, maximum displacement from equilibrium called amplitude, and spatial extension known as wavelength. There are numerous physical systems whose motion, once initiated, exemplifies a wave motion. Among these are the motion of a bob in a simple pendulum, and the motion of a mass, attached at one end to an ideal spring and with its other end fastened to a rigid support (Chapter 9). The mass attached to a hanging spring, if slightly pulled downward and released, executes a periodic motion about its equilibrium position. The propagation of sound is another example of a wave motion as is the propagation of light as electromagnetic waves. This chapter presents a mathematical description of waves, a discussion of their main properties, and establishes a connection between the mathematical model of a wave and its real properties.

11.1 WAVE MOTION

Waves can be generalized in many forms: square, sawtooth, sinusoidal, and so on. A wave is considered ideal if it consists of identical oscillations propagating in space and time, and regardless of their origin and form, they have basic common features. They carry energy and momentum. From a mathematical perspective, they are described by a sinusoidal function, sine or cosine, which are simple functions that facilitate calculations of fundamental physical properties of waves. The periodic aspect of waves in space and time is illustrated in Figures 11.1 and 11.2, respectively. Figure 11.1 demonstrates the periodic change of the disturbance y at various values of x taken at a fixed instant. This can be likened to a camera snapshot. However, the plot in Figure 11.2 depicts the variation of the disturbance y versus time t as observed at a fixed position x. The following is a brief description of the common features of waves.

Amplitude, A: Is the maximum displacement of the wave above or below the position of equilibrium. The equilibrium position is the middle position, or level, about which the system is oscillating in a periodic manner. For a wave described by the following sine equation:

$$y = A \sin(kx - \omega t), \tag{11.1}$$

where A is the amplitude of the wave, and the argument $(kx - \omega t)$ is its phase.

Intensity, I: The energy intensity of a wave, that is, energy perunit area per unit time is proprtional to the square of the absolute value of its ampliude. Thus for the wave, described above, the intensity I is

$$I \propto |A|^2. \tag{11.2}$$

The intensity of a wave will be discussed separately in Section 11.4.

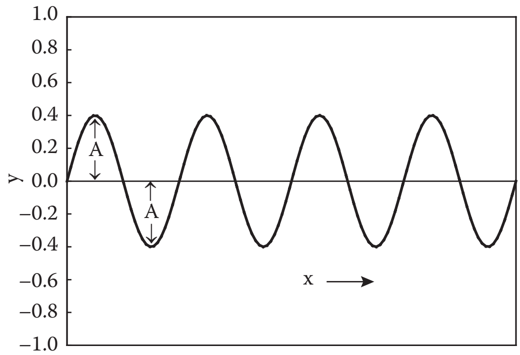


FIGURE 11.1 Representation of a sine wave as a function of x at a fixed instant.

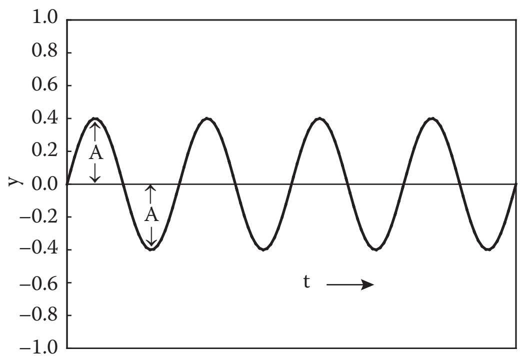


FIGURE 11.2 Representation of a sine wave as a function of time t at a fixed position x.

Wavelength, λ: This is the distance along the wave between two nearest identical points. For example, it is the distance between two successive peaks or troughs. Two points on the wave that are identical are said to be of the same phase, that is, oscillating in exactly the same manner. In Figure 11.3, points c_1 and c_2 are of one phase, and b_1 and b_2 are of another phase. However, points c_2 and b_2 are of different phases, because they tend to oscillate in different directions at any instant of time.

Period, T: The period T of a wave is the time it takes a point on the wave to execute one full cycle. The point would have traveled a distance equal to λ.

Linear frequency, f: This is the number of cycles generated per second. Its units are s^{-1}, or Hertz (Hz). It is equal to the inverse of the period T of the wave, that is,

$$f = \frac{1}{T}. \tag{11.3}$$

Angular frequency, ω: This is related to the linear frequency as follows

$$\omega = 2\pi sf. \tag{11.4}$$

Or using Equation 11.3

$$\omega = 2\pi\frac{1}{T}. \tag{11.5}$$

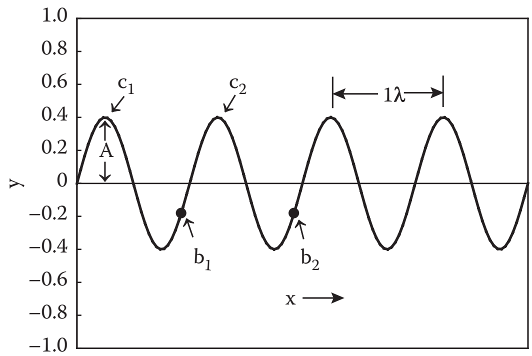


FIGURE 11.3 Demonstration of the notion of phase showing that points c_1 and c_2 are of identical phase and points b_1 and b_2 are of another identical phase.

Wave number, k: This is also called propagation number, and is defined as the number of waves per unit length. It is given by

$$k = \frac{2\pi}{\lambda}. \tag{11.6}$$

Velocity, v: The rate at which the disturbance propagates defines the velocity of the wave, and this is given by

$$v = \frac{\lambda}{T}. \tag{11.7}$$

As can be noticed, v is equal to a distance of one wavelength traveled by a point on the wave divided by its period, T. Using Equation 11.3, the velocity in Equation 11.7 is

$$v = \lambda f. \tag{11.8}$$

11.2 TYPES OF WAVES

Waves, regardless of their origin, are categorized into two kinds. In one kind, the displacement of the constituents of the medium in which it is traveling is along the direction of propagation of the wave. These are known as longitudinal waves, of which sound waves are a typical example. In the second kind, the displacement of the medium is perpendicular to the direction of propagation of the wave. These are called transverse waves. Waves produced on a string are a typical example of this type. Points on these waves oscillate in directions perpendicular to their direction of propagation.

As for longitudinal waves such as sound, their velocity depends on the medium of propagation. In case the medium is a solid, Young's modulus and the density of the solid are important factors. However, propagation of these waves in a gas depends on the specific heat of the gas, its pressure, and density. The speed of sound in air at 20°C is 343.37 m/s (CRC 70th Edition, 1989–1990, p. E-49). As for transverse waves, their velocity also depends on the medium of propagation. For electromagnetic waves, the speed of light in vacuum is constant ($c = 3.00 \times 10^8$ m/s). However, in any other transparent or translucent medium, its speed v is less than c. As electromagnetic waves are treated elaborately in later chapters, the current treatment will be limited to basic properties of waves.

EXAMPLE 11.1

The wavelegth of a wave is 4.00 m. Find the speed of this wave if its frequency is 5.00 Hz.

Solution

From Equation 11.8,

$$v = \lambda f.$$

Thus,

$$v = (4.00\text{ m})\ (5.00\text{ s}^{-1)}) = 20.0\text{ m/s}.$$

Example 11.2

Consider a wave whose frequency is doubled while its speed remained constant. Describe quantitatively the changes that will occur to the following:

a. The wave's period
b. Its wavelength

Solution

a. Referring to quantities before and after the frequency is doubled by subscripts 1 and 2, then

$$f_2 = 2f_1.$$

Using Equation 11.3

$$\frac{1}{T_2} = 2\left(\frac{1}{T_1}\right).$$

This reduces to

$$T_2 = \frac{T_1}{2}. \tag{11.9}$$

b. From Equation 11.8

$$v_1 = \lambda_1 f_1 \tag{11.10a}$$

and

$$v_2 = \lambda_2 f_2, \tag{11.10b}$$

and $v_1 = v_2$ since the velocity does not depend on the wave's frequency. Dividing Equation 11.10a by 11.10b, and substituting for

$$f_2 = 2f_1,$$

gives

$$1 = \frac{\lambda_1 f_1}{\lambda_2 (2f_1)}.$$

This reduces to

$$\lambda_2 = \frac{\lambda_1}{2}. \tag{11.11}$$

Analysis

It can be noticed from Equations 11.9 and 11.11 that doubling the frequency of a wave while keeping its velocity constant results in its period reduced to half, and its wavelength also reduced to 1/2.

11.3 MATHEMATICAL TREATMENT OF AN IDEAL WAVE

Figure 11.4 shows two sinusoidal waves: (a) sine wave and (b) cosine wave, each a function of position, both depicted at a certain fixed instant. The direction of propagation is along the x axis. Except for a displacement along the x axis by one-fourth wavelength for (b) over (a), the cosine wave is identical to the sine wave. That is why both functions are equally useful in describing a wave. Considering the sine form with both its position and time dependences included, the wave function becomes

$$y = A\sin(kx - \omega t), \tag{11.12}$$

where y is the disturbance along y, and x is the direction of propagation.

In reference to a right angle triangle, the sine function is defined as the ratio between one of the sides of a right angle triangle and the hypotenuse. Therefore, neither the sine nor the cosine function has units. If x is in meters and t in seconds, the quantities k (= $2\pi/\lambda$) and ω (= $2\pi(1/T)$) make Equation 11.12 take the form

$$y = A\sin 2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right). \tag{11.13}$$

Equations 11.12 and 11.13 both show that for $x = 0$ and $t = 0$, the disturbance is $y = 0$. If it happens that a wave starts from a position $x \neq 0$, y then will have a finite value, and from Equation 11.13, the sine function at $t = 0$ will have a nonzero value. In such a case, an angle θ is added to the argument in both Equations 11.12 and 11.13. Theta is called the initial phase or phase constant of the wave. Equation 11.12 then becomes

$$y = A\sin(kx - \omega t + \theta) \tag{11.14}$$

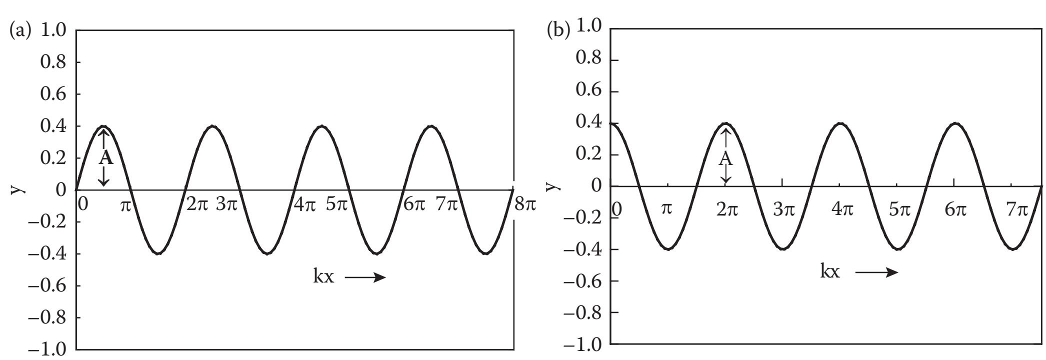


FIGURE 11.4 Two sinusoidal waves: (a) sine and (b) cosine, each of which is a function of position representing a snapshot at a certain instant that is fixed.

or

$$y = A \sin\left[\left\{2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right)\right\} + \theta\right]; \tag{11.15}$$

the angle θ is measured in radians.

11.3.1 Special Remarks

1. The difference between Equations 11.14 and 11.12 can be described as a distance by which the wave in Equation 11.14 is ahead of that expressed by Equation 11.12 when the two are observed at the same fixed instant, which for simplicity may be set to zero. This distance is also called path dfference and is given by

$$x = \frac{\theta}{2\pi}\lambda. \tag{11.16}$$

As from Equation 11.6, $k = 2\pi/\lambda$, the above equation can also have the form

$$x = \frac{\theta}{k} = \left(\frac{1}{k}\right)\theta. \tag{11.17}$$

That is, (1/k) serves as a converting factor from an angular phase difference to a path difference.

2. The difference between Equations 11.14 and 11.12 can be described as a time interval by which the wave in Equation 11.14 is ahead of that expressed by Equation 11.12 when the two waves are monitored at the same position, $x = 0$. This interval is called the temporal dfference and is given by

$$t_{\text{temp diff}} = \frac{\theta}{2\pi}T. \tag{11.18}$$

As $\omega = 2\pi(1/T)$, the above equation can also be reexpressed in the form

$$t_{\text{temp diff}} = \left(\frac{1}{\omega}\right)\theta. \tag{11.19}$$

That is, (1/ω) serves as a converting factor from angular phase difference to temporal difference.

EXAMPLE 11.3

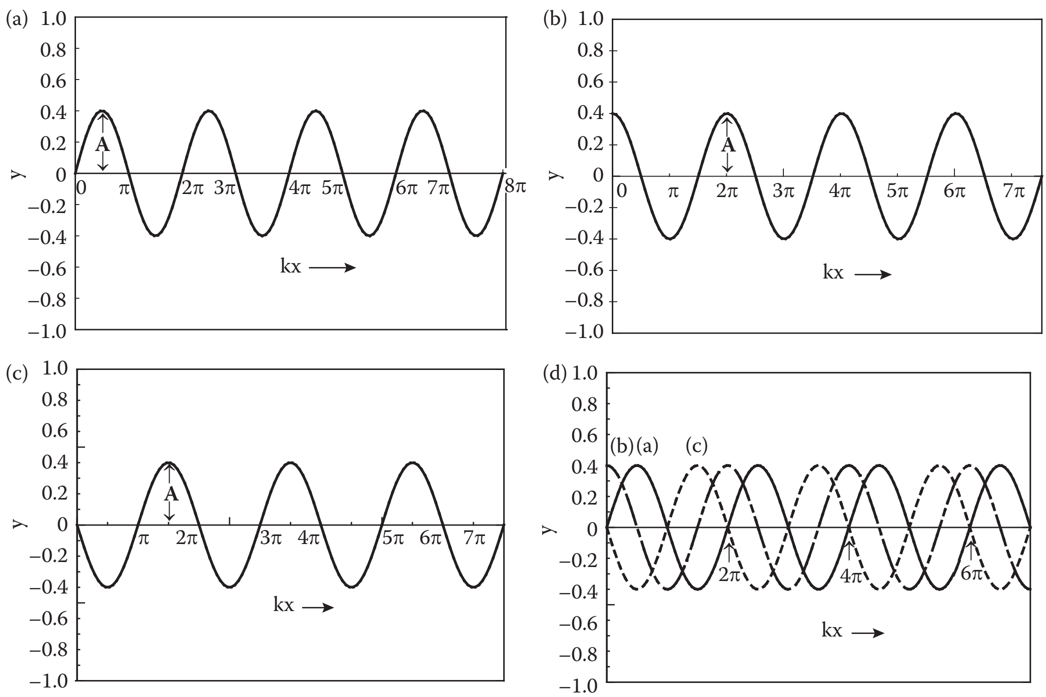
Using Equation 11.14, consider the following values for the initial phase: (a) $\theta = 0$, (b) $\theta = \pi/2$, and (c) $\theta = \pi$. In each case make a sketch of the wave form, and compare it with the wave form, $y = A \sin(kx - \omega t)$. Comment on what is common and what is different between the two wave forms.

Solution

To simplify the solution and discussion, setting $t = 0$ would be the most convenient. The following figure (part (a–d)) are four plots that correspond to the following:

a. The wave form given by Equation 11.12
b. The wave form that describes case (b)

c. The wave form that describes case (c)
d. The above three plots, all in one diagram, offer a direct comparison between them
   a. Plot of Equation 11.14: $y = A \sin(kx - \omega t)$
   b. $y = A \sin(kx - \omega t + \pi/2)$
   c. $y = A \sin(kx - \omega t + \pi)$
   d. The above three cases combined



Three sine waves of different initial phases: (a) $y = A \sin(kx - \omega t)$, (b) $y = A \sin(kx - \omega t + \pi/2)$, (c) $y = A \sin(kx - \omega t + \pi)$, and (d) all three cases combined.

Analysis

1. It can be noticed that wave form (c) is the negative of wave form (a); and (a) and (c) are mirror image of each other with mirror aligned along the x axis. This is due to the π (= 180°) angular phase difference between the two waves. Such waves are called out of phase. However, except for a displacement of $\pi/2$ (= 90°) between (a) and (b), they are identical. Wave (b) is ahead of (a) by $\pi/2$. If wave (b) is diplaced to the right by one-fourth of a cycle, it becomes identical to (a).
2. The waves in the three cases are identical in amplitude, wavelength, frequency, and speed.

EXAMPLE 11.4

The following: $y_1 = 0.20 \sin(0.22\pi x)$ and $y_2 = 0.40 \sin(0.22\pi x + 0.44)$ are the wave equations of two waves at $t = 0.00$ s. If x and y are in cm, determine

a. The phase difference between the two waves
b. The wave number of each of the two waves
c. The path difference between the two waves

Solution

a. The phase difference θ between the two waves is 0.44 rad.
b. From either equations, the wave number is

$$k = 0.22\pi \text{ cm}^{-1}.$$

c. As

$$k = \frac{2\pi}{\lambda},$$

then

$$0.22\pi = \frac{2\pi}{\lambda},$$

$$\lambda = 9.1 \text{ cm}.$$

As every complete cycle (i.e., one wavelength) of the wave corresponds to an angle of 2π, then the distance along the x axis that corresponds to an angular phase difference of 0.44 rad is

$$x = \left(\frac{0.44 \text{ rad}}{2\pi \text{ rad}}\right)(9.1 \text{ cm}) = 0.64 \text{ cm}.$$

Analysis

The answer shows that the path difference is so small compared to the wavelength of either wave. It is only 0.07 of the wavelength. If the angular phase difference (0.44 rad) is compared to 2π that represents a full cycle (= one wavelength), one also finds that it is 0.07 as it should be.

11.4 INTENSITY OF WAVES

The energy intensity of a wave is directly proprtional to its energy; hence it must always be positive. Figure 11.5 shows a sine wave of amplitude A, plot a. The square of the wave is shown in plot b. The intensity of the wave is determined from the absolute square, A^2.

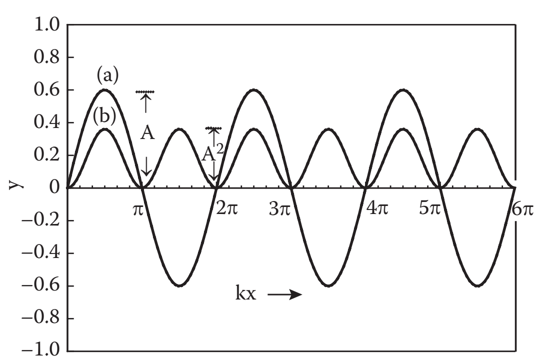


FIGURE 11.5 Depiction of two plots: (a) a sine wave of amplitude A and (b) the wave squared showing its intensity $I \propto A^2$.

EXAMPLE 11.5

Refer to Example 11.4 and determine the ratio between the intensities of the two waves described in the example.

Solution

The two waves y_1 and y_2 in the previous example have the amplitudes $A_1 = 0.20$ cm and $A_2 = 0.40$ cm, respectively.

As the intensity of a wave is proportional to the absolute square value of its amplitude, then

$$\frac{I_1}{I_2} = \frac{(A_1)^2}{(A_2)^2} = \frac{(0.20\ \text{cm})^2}{(0.40\ \text{cm})^2} = 0.25. \tag{11.20}$$

Analysis

As the intensity in a wave is indicative of the energy per unit area per second, then a unit area of a surface, a distance r away from both waves, will be crossed each second by energy from the second wave equal to four times the energy from the first one.

11.5 SUPERPOSITION OF WAVES: INTERFERENCE AND STANDING WAVES

The superposition principle applies to all kinds of ideal periodic waves, transverse and longitudinal, as long as the medium has the same physical properties along all directions. In such a case, the medium is described as isotropic or homogeneous. Through the principle of superposition, the resultant of two or more waves that are simultaneously overlapping throughout their propagation can be determined. The waves interferring with each other may have equal or different amplitudes, frequencies, and constant or random phase differences. The simplest interference may best be described for waves emitted by two sources. Depending on the amplitudes and frequencies of the two waves and on the phase difference between them at a receiving end, two cases are of special interest. These are when the resultant wave is an enforcement or suppression of each other. The first type is described as totally constructive, while the latter as totally destructive interference. Both types of interference are employed in the design of optical filters, special mirrors, and lenses for various applications. Another special type of superposition produces standing waves. These are produced by two waves, transverse or longitudinal, of equal amplitude, same frequency, and of 180° phase difference, propagating in opposite directions. While a thorough treatment of interference between light waves is left for Chapter 21, study of standing waves is discussed below.

Figure 11.6 illustrates the situation for two waves, incident traveling wave, I, and reflected one, R, both of equal ampliudes. The resultant of the two waves at any point is the algebraic sum of the

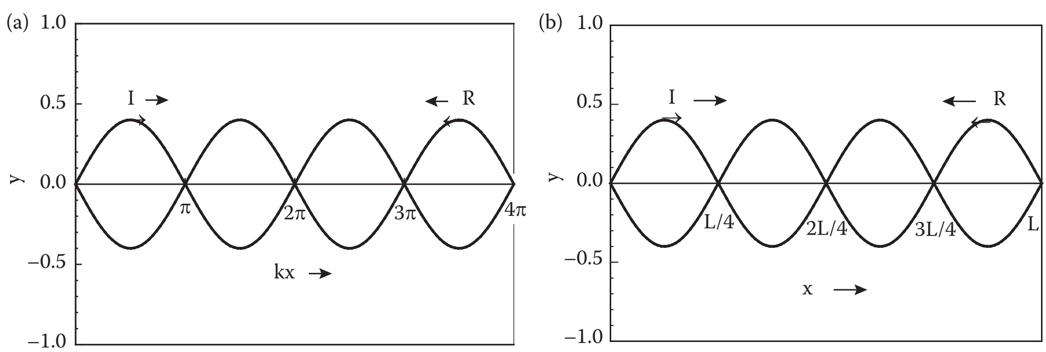


FIGURE 11.6 Standing waves: (a) y versus kx displaying the nodes in terms of π and (b) y versus x displaying the nodes along the string at equal distances from each other.

disturbances produced by them at that point. In this case, superposition produces standing waves with various modes of vibrations that depend on the tension applied to the stretched string. Points where the resultant disturbance is zero are called nodes, and those where the resultant is maximmum on either side of the node are called antinodes. The envelope engulfing the space between two successive nodes is called a segment. In the figure, the sketch displays four segments along a string of length L. Any node is one-half wavelength away from the following node, and in any of the generated modes of vibrations, there is no net transfer of energy. Therefore, the notion of calling standing waves is a misnomer. Two plots are chosen here, one (Figure 11.6a) y versus kx displaying the nodes in terms of π, which corresponds to half of the cycle of each of the two superimposed waves. However, a plot of y versus x would display the nodes along the string equally distant from each other (Figure 11.6b).

The wavelengths λ_n of the standing waves are related to the string length L via the following relation

$$\lambda_n = \frac{2L}{n}, \tag{11.21}$$

where n is an integer, n = 1, 2, 3,. . .. The lowest value for the frequency f (n = 1) is called the fundamental frequency or the first harmonic. As Equation 11.21 shows, the wavelength will be

$$\lambda_1 = 2\,L.$$

From Equation 11.8, the frequencies that asssociate the wavelengths λ_n are

$$f_n = \frac{v}{\lambda_n},$$

or

$$f_n = n\left(\frac{v}{2L}\right). \tag{11.22}$$

As $f_1 = (v/2\,L)$, f_n turns out to be a multiple of the fundamental frequency f_1. That is,

$$f_n = nf_1, \quad n = 1, 2, 3,\ldots. \tag{11.23}$$

Figure 11.7 demonstrates the first harmonic (n = 1), the mode that has two nodes, one at each end of the string. For this mode, there is one envelope.

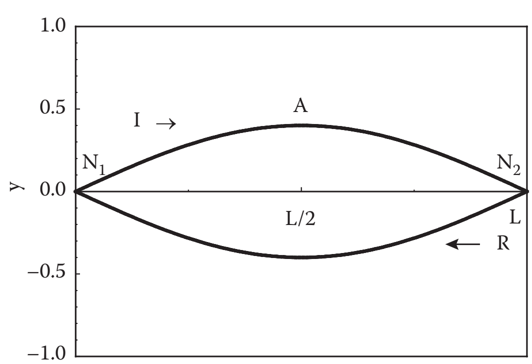


FIGURE 11.7 Depiction of the first harmonic (n = 1). The mode has two nodes and one envelope.

Modes of standing sound waves produced in an air column in a glass tube closed at one end and open at other end are in many aspects similar to those produced in a string. The basic difference resides in the wavelength of the first harmonic being four times the length of the air column. Figure 11.8 shows the first and third harmonics of a standing sound wave. In (a), the length of the air column is one-fourth the wavelength, while in (b) the length of the air column is three-fourths the wavelength, giving

$$f_1 = \frac{v}{4L},$$

f_n turns out to be a multiple of the fundamental frequency f_1. That is, for the one end closed air column,

$$f_n = nf_1, \quad n = 1, 3, 5, \ldots \tag{11.24}$$

Again, the lowest value for f (n = 1) is called the fundamental frequency or the first harmonic; n = 3 is the third harmonic, n = 5 is the fifth harmonic, and so forth.

The general form for the wavelengths λ_n of the standing sound waves relates to the length L of the air column as follows:

$$\lambda_n = \frac{4L}{n}, \tag{11.25}$$

where n is an integer, n = 1, 3, 5,. . .. As Equation 11.25 shows, the wavelength will be

$$\lambda_1 = 4\,L.$$

From Equation 11.8, the frequencies that asssociate the wavelengths λ_n are

$$f_n = \frac{v}{\lambda_n},$$

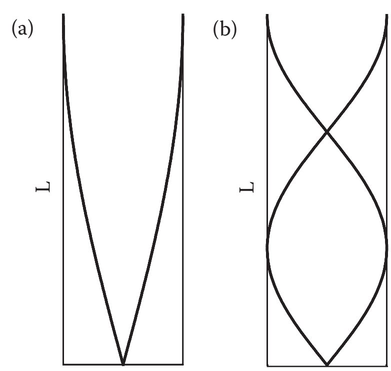


FIGURE 11.8 Sketch of the first and third harmonics of a standing sound wave produced in an air column open from one side. (a) The length of the air column is one-fourth the wavelength. (b) The length of the air column is three-fourths the wavelength.

or

$$f_n = n\left(\frac{v}{4L}\right). \tag{11.26}$$

The discussion can be extended to a glass tube open on both ends. Figure 11.9 shows the lowest two harmonics in which we can notice that the first harmonic (a) has a wavelength equal to two times the tube's length, while the next harmonic (b) has a wavelength equal to the length of the tube.

The general form for the wavelengths λ_n of the standing sound waves relates to the length L of the air column in the open-ended tube as follows:

$$\lambda_n = \frac{2L}{n}, \tag{11.27}$$

where n is an integer, n = 1, 2, 3, …. As Equation 11.27 shows, the wavelength will be

$$\lambda_1 = 2L.$$

From Equation 11.8, the frequencies that associate the wavelengths λ_n are

$$f_n = \frac{v}{\lambda_n}$$

and

$$f_n = n\left(\frac{v}{2L}\right), \tag{11.28}$$

giving

$$f_1 = \frac{v}{2L},$$

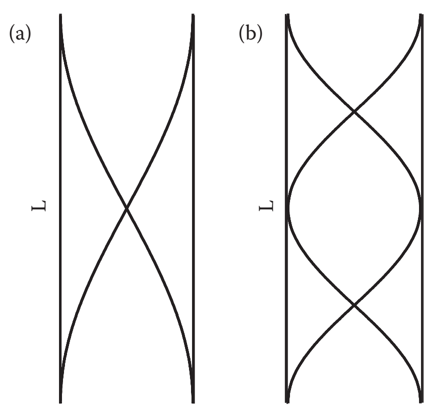


FIGURE 11.9 The lowest two harmonics of a standing sound wave produced in an air column of length L open from two sides. (a) L = λ/2 and (b) L = λ.

f_n turns out to be a multiple of the fundamental frequency f_1. That is,

$$f_n = nf_1, \quad n = 1, 2, 3, \ldots \tag{11.29}$$

Equation 11.29 demonstrates the parallel between the nth harmonic of the standing waves produced in a open-ended glass and those produced on a string (Equation 11.23).

EXAMPLE 11.6

A string 90.0 cm long is set into vibration at one end while its other end is fixed. For the second harmonic ($n = 2$) wave that is reflected at the string's fixed end producing a standing wave, find

a. The distance between two nodes of the standing wave
b. The distance between two nodes of the standing wave for the third harmonic

Solution

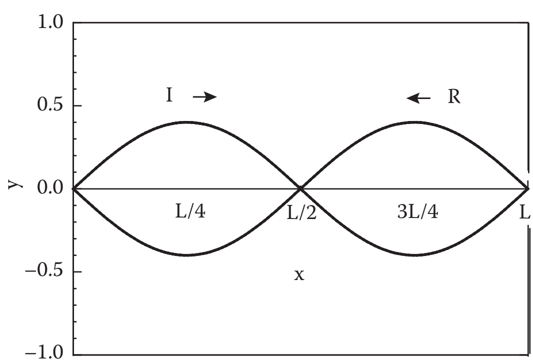
a. From Equation 11.21,

$$\lambda_n = \frac{2L}{n}.$$

Thus the second harmonic, that is, $n = 2$, gives

$$\lambda_2 = L = 90.0 \text{ cm}.$$

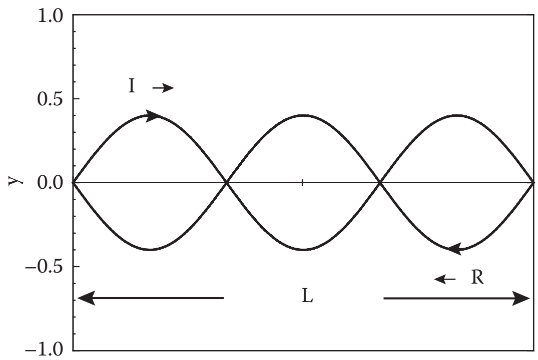
This means that the wavelength of the original wave propagating along the string is 90.0 cm. The wave reflected from the other end of the string will also have a wavelength of 90.0 cm. These two waves intersect at the middle of the string producing two envelopes (the figure below). There will be three nodes, one at each end and a third midway between the two ends. The distance between two nodes is then 45.0 cm.



b. For the third harmonic, $n = 3$ gives the wavelength

$$\lambda_3 = \frac{2L}{3} = 60.0 \text{ cm}.$$

For this harmonic there are two waves, one propagating in one direction, and the other reflected in the opposite direction. These intersect along the string twice producing three envelopes (the following figure), giving four nodes, two at the two ends of the string and two others at equal distances from the ends. Thus, the distance between two nodes will be 30.0 cm.



Analysis

1. Following the same method, the distance between the nodes can be obtained for any harmonic generated on the string.
2. From the proceeding treatment, it can be noticed that the number of the nodes, N, for any harmonic, n, generated on a given string follows the simple rule $N = n + 1$.

PROBLEMS

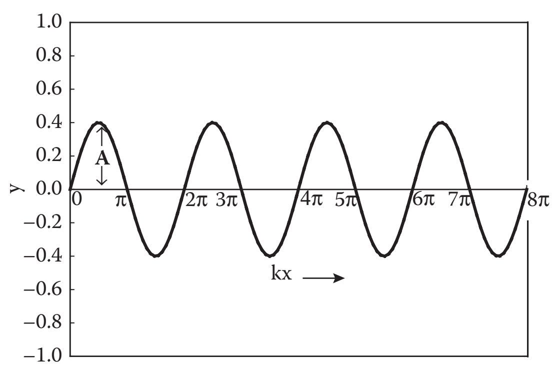
11.1 An oscillator of 120.0 Hz, driving a pulse on a string, generates a wave of 50.0 cm wavelength and 8.0 cm amplitude. Determine the period of the wave and its speed.

11.2 Calculate the frequency range of an electromagnetic wave that corresponds to a wavelength range 20.0 m – 2.00×10^{-7} m. (*Hint*: The speed of light is 3.00×10^8 m/s.)

11.3 Consider a wave whose angular frequency is $\omega = 0.600$ rad/s and wave number is $k = 6.00 \times 10^{-3}\ m^{-1}$. Calculate the wavelength and speed of the wave.

11.4 Consider the wave depicted in the figure below that is generated along a 1.00 m string by an oscillator vibrating at 50.0 Hz; x and y are in meters. For this wave determine

a. The wavelength
b. The period
c. The speed



11.5 Refer to the diagram used in the previous problem, and generate a new version of the plot such that the horizontal axis is merely in terms of x only in place of kx. Use the new plot to determine the wavelength of the wave.

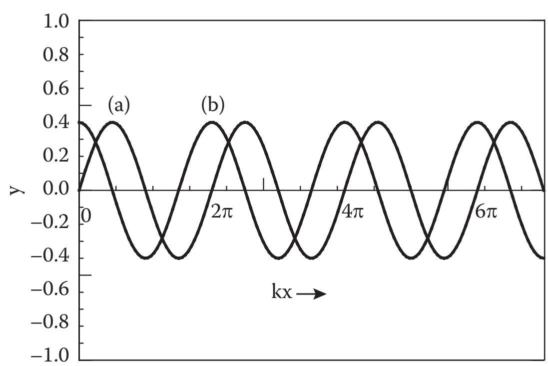
11.6 A wave described by the equation $y = 0.20 \sin(4.00\pi x - 12\pi t + 0.33)$ is set up on a long string. If x and y are in cm, and t in seconds, determine
a. The wave number of the wave
b. The wavelength of the wave
c. The speed of the wave

11.7 For the wave in the previous problem, determine
a. The transvers displacement of the wave at $x = 2.00$ cm and $t = 0.200$ s
b. The initial phase angle of the wave in radians at $x = 2.00$ cm and $t = 0.200$ s

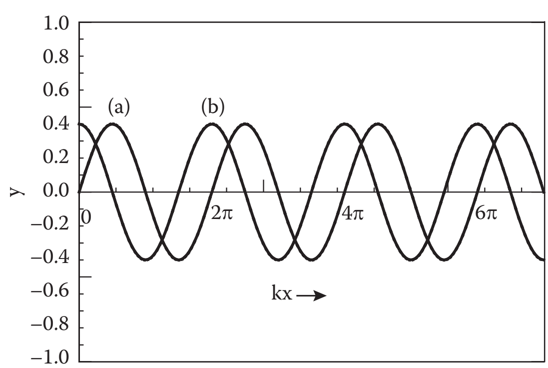
11.8 Sketch the following function, $y = A \cos(kx - \omega t)$, and compare it with the sketch in Example 11.3a.

11.9 The following: $y_1 = 0.20 \sin(2.0\pi x - \omega t)$ and $y_2 = 0.80 \cos(2.0\pi x - \omega t + 0.44)$ are the wave equations of two waves at $t = 0.0$ s. Knowing that for any angle θ, $\cos(90° - \theta) = \sin \theta$, and if x and y are in m, determine
a. The phase difference between the two waves
b. The wave number of each of the two waves
c. Compare the intensitiy of first wave and the second one by finding the ratio between them

11.10 The figure below shows a snapshot of a wave (a) at time $t = 0$ and (b) at time $t = 0.104$ s. The distance x in the diagram is in cm. Using the graph, determine
a. The wavelength of the wave
b. The speed of the wave

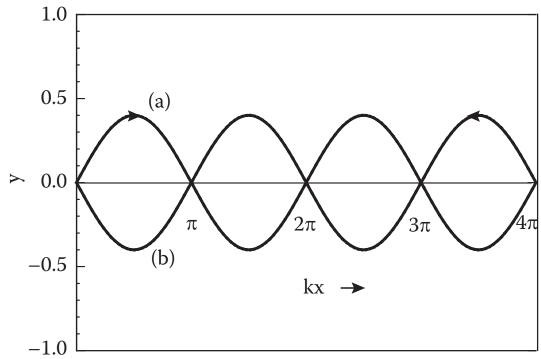


11.11 The figure below shows a snapshot of two waves at a certain instant $t = 0$. Apply the superposition principle to determine the resultant of the two waves by adding algebraically their disturbances at the following selected values: $(kx) = \pi/2$, $(kx) = \pi$, $(kx) = 3\pi/2$. Add to the figure a plot of the resultant.

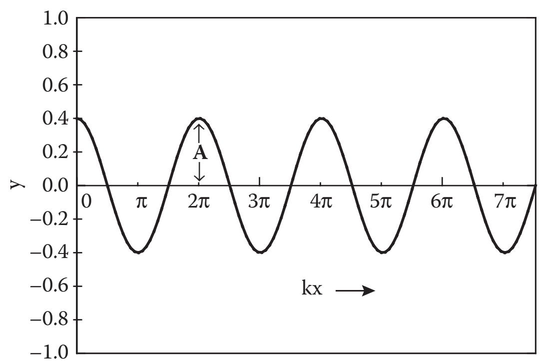


11.12 The figure below shows a 1.00 m string vibrating in four segments to a frequency set to 368 Hz.

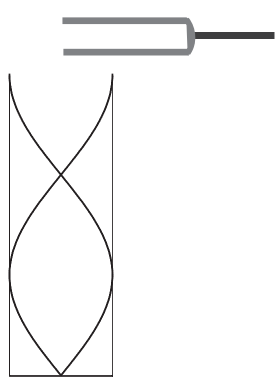
a. Determine the fundamental frequency.
b. Determine the velocity of the wave.
c. The frequency that would make the string oscillate in six segments.



11.13 Consider the wave depicted in the figure below and make a plot of its mirror wave, that is, its reflection through the x axis. Suggest the mathematical wave form that the mirror wave should have.



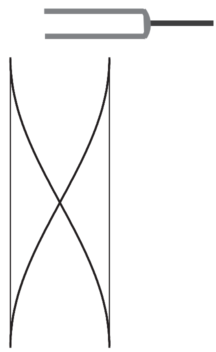
11.14 A vertically held glass tube 1.00 m long, closed at the bottom side and open at the top, and a tuning fork oscillating at frequency of 256 c/s were used to produce the resonance depicted in the figure below. Find the speed of air in the tube.



11.15 A vertically held glass tube 0.680 m long, open from both ends, and a tuning fork oscillating at frequency of 256 c/s were used to produce the first harmonic (the figure below).

a. Find the speed of air in the tube.

b. What would be the wavelength of the fourth harmonic?



12 Fluids

Liquids and gases have many common interesting properties. Among these is the absence of a rigid shape or form. Both take the shape of the enclosure they fill. Buoyancy is another common property, and is addressed by Archimedes' principle. They also follow many common rules. This chapter addresses some of these properties, and the formulation of the mechanics of fluids, using Newton's laws.

12.1 PRESSURE: DEFINITION AND UNITS

Pressure exerted on a surface is produced when a force acts on it. It is defined as the force acting normal to an element of the surface divided by the element area A. That is,

$$P = \frac{F}{A}. \tag{12.1}$$

The SI unit for pressure is the Pascal (Pa). That is,

$$1.0\ \text{Pa} = 1.0\ \text{N/m}^2.$$

The unit is attributed to the French physicist Blaise Pascal (1623–1662). Other units in use are the "bar" and the atmospheric pressure, "atm." These are related to the Pascal as follows:

$$1\ \text{bar} = 10^5\ \text{Pa},$$

$$1.00\ \text{atm} = 1.01 \times 10^5\ \text{Pa}. \tag{12.2}$$

Atmospheric pressure results from air molecules acting with a force perpendicular to an area. As the element of area at a particular location can be at any orientation, the pressure at that location is the same in all directions. It also can be noted from Equation 12.2 that from a practical perspective, the "Pascal" is a small unit. Therefore, other units such as the bar are often used in vacuum systems and in designing vacuum gauges. Another unit, the British unit, lb/in^2, is related to the Pascal as follows

$$1.00\ \text{Pa} = 1.46 \times 10^{-4}\ \text{lb/in}^2.$$

Thus, atmospheric pressure when expressed in lb/in^2 has the value

$$1.00\ \text{atm} = 14.7\ \text{lb/in}^2. \tag{12.3}$$

EXAMPLE 12.1

Find the pressure exerted on the ground by each foot of an 822 kg elephant asssuming that each of his feet measures 0.0420 m^2 and that the weight is uniformly distributed on his feet.

Solution

From Equation 12.1,

$$P = \frac{F}{A},$$

where F is simply the elephant's weight, W. That is,

$$F = W = (822\ \text{kg})\ (9.80\ \text{m/s}^2) = 8060\ \text{N}.$$

Thus, the pressure P_w on the ground from the elephant's foot resulting from his weight becomes

$$P_w = \frac{8060\ \text{N}}{4 \times (0.0420\ \text{m}^2)} = 4.80 \times 10^4 \text{N/m}^2.$$

Comparing this value with the atmospheric pressure makes $P_w = 0.475$ atm.

Analysis

Adding 1.0 atm to the above value gives the total pressure from each foot. The total pressure, P, is usually referred to as the absolute pressure.

The above example draws special attention to another more practical quantity, known as gauge pressure. This is denoted by P_g. Gauge pressure in tires of bicycles, cars, airoplanes, and vacuum systems is among its applications. P_g at any location is defined as the difference between the absolute value of pressure at that location minus the atmospheric pressure. That is,

$$P_g = P - P_{atm} \tag{12.4a}$$

or

$$P = P_g + P_{atm}. \tag{12.4b}$$

EXAMPLE 12.2

Knowing that the absolute pressure on a car tire is 3.20×10^5 Pa, find the gauge pressure inside the tire, expressing it in SI and British units.

Solution

From Equation 12.4a, the gauge pressure is

$$P_g = P - P_{atm}.$$

Therefore,

$$P_g = 3.20 \times 10^5\ \text{Pa} - 1.01 \times 10^5\ \text{Pa} = 2.19 \times 10^5\ \text{Pa}.$$

In British units, P_g becomes

$$P_g = 2.19 \times 10^5\ \text{Pa} \times (1.46 \times 10^{-4}\ \text{lb/in}^2/\text{Pa}) = 32.0\ \text{lb/in}^2.$$

Analysis

Usually, air pumps are calibrated in British units rather than in Pa. Hence, the gauge pressure in British units would seem more common for the pressure in the tire.

12.2 PRESSURE WITHIN FLUIDS

In contrast to solids, a fluid that is in equilibrium within a container has an interesting property with regard to its response to forces acting on it. A finite external force acting on a gas or a liquid may result in changing its shape or volume, but the pressure that such an external force creates is transmitted uniformly throughout the whole fluid. Thus, the pressure at a certain depth h within a fluid has the same value, P_h acting in all directions (Figure 12.1). In addition, a fluid in a container exerts forces that are perpendicular to the container surfaces. Otherwise, components of fluid forces that are parallel to the surfaces would, from Newton's third law, subject the fluid molecules to equal reaction forces in the opposite direction. That then would result in moving these molecules, which is contrary to observations in a static fluid.

12.3 DENSITY OF FLUIDS

Density is another physical quantity of importance in formulating the mechanics of fluids. The pressure at a given location in a fluid depends on the mass of the column of the fluid above that location. The density ρ of an object is defined as the mass per unit volume. In symbols,

$$\rho = \frac{m}{V}. \tag{12.5}$$

The SI unit of density is kg/m^3. Table 12.1 lists the densities of some fluids.

EXAMPLE 12.3

Knowing that the density of gold is 19.3×10^3 kg/m^3, determine the volume of

a. One ounce (0.028 kg) of gold
b. One pound (0.454 kg) of gold

Solution

a. From Equation 12.5,

$$\rho = \frac{m}{V}.$$

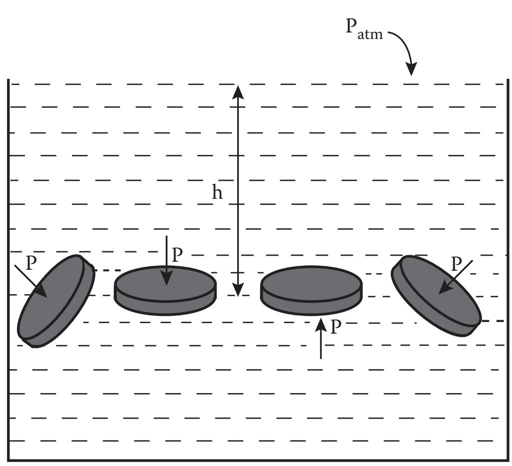


FIGURE 12.1 Several identical area elements in a liquid, all at the same depth, but of different orientations. The pressure on all elements is the same.

TABLE 12.1
Densities of Several Fluids

Fluid	Density, ρ (kg/m^3)
Water	1.0×10^3
Benzene	0.899×10^3
Alcohol	0.79×10^3
Mercury	13.6×10^3
Air	1.20
Oxygen	1.43
Hydrogen	8.99×10^{-2}
Helium	1.79×10^{-1}
Milk	$1.03–1.04 \times 10^3$
Olive oil	0.918×10^3
Glycerin	1.26×10^3

Source: *CRC Handbook of Chemistry and Physics*, 70th Edition, 1989–1990, pp. B-19, B-20, B-27, F-3, F-10. With permission.

Thus the volume of a mass m is

$$V = \frac{m}{\rho}.$$

Upon substituting for the given values, the volume of one ounce of gold becomes:

$$V = \frac{0.028 \text{ kg}}{19.3 \times 10^3 \text{kg/m}^3} = 1.5 \times 10^{-6} \text{m}^3.$$

In the cgs system of units, the above volume is 1.5 cm^3.

b. Since one pound equals 16 ounces, the volume of one pound, rounded to two significant figures, becomes

$$V = 24 \times 10^{-6} \text{ m}^3 = 24 \text{ cm}^3.$$

Analysis

An ounce of impure gold would naturally have a density different from the density of pure gold, and, hence for a given mass, would have a different volume. Thus, a careful measure of the volume of such an ounce would reveal it as a fake.

12.4 VARIATION OF PRESSURE WITH DEPTH IN A STATIC FLUID

This property is experienced in oceans, rivers, lakes, and contained liquid. Pressure in a lake at a certain depth beneath its surface is greater at lower depths. This is due to the larger weight of the water column above the deeper locations compared to those nearer to the surface.

Consider a static liquid of density ρ confined to a container (Figure 12.2). Consider an infinitesimal area within the liquid in the form of a visualized horizontal disk of circular area A, thickness Δy, mass Δm, and volume ΔV at a depth y below the surface of the water, where the origin O is conveniently placed; y is positively downward. Let the pressure on the upper side of the disk be P, while on its lower side be $P + \Delta P$.

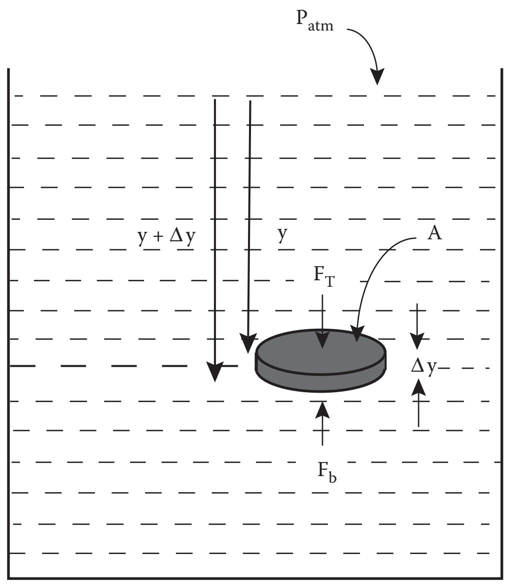


FIGURE 12.2 A liquid confined to a container. The shaded area is a visualized infinitesimal horizontal disk of the liquid of top surface area A located at depth y below the surface.

From the geometry, the bottom of the disk is at depth y + Δy, and since the liquid is in a static state, the resultant force on the disk is zero. Assume that the force on the lower surface of the disk is F_b, and on its top surface F_T, then from Newton's second law applied to the mass Δm,

$$\sum_{\text{vert.}} F_{\text{disc}} = F_T + (\Delta m)(g) - F_b = 0.$$

Using Equation 12.1 for F_T and F_B, the above equation becomes

$$(P)\,A + (\Delta m)(g) - (P + \Delta P)\,A = 0.$$

Canceling identical terms reduces the above equation to

$$-\Delta P\,A + (\Delta m)(g) = 0.$$

Substituting for $\Delta m = \rho\Delta V$ gives

$$(\Delta P)\,A = (\rho\Delta V)(g).$$

Since $\Delta V = A\Delta y$, then

$$(\Delta P)\,A + (\rho A\,\Delta y)(g) = 0.$$

That is,

$$(\Delta P) = (\rho g)\Delta y. \tag{12.6}$$

Equation 12.6 may also be written as

$$\frac{\Delta P}{\Delta y} = \rho g. \tag{12.7}$$

Notice that $\Delta P/\Delta y$ in Equation 12.7 is as follows:

1. Linear in ρ, which means that it is directly proportional to ρ.
2. Positive, which means that the pressure increases as y beneath the water surface increases. Thus as expected, the pressure becomes smaller at locations that are closer to the surface.

From Equation 12.6, the difference in pressure between two locations that are at depths y_1 and y_2 from the surface is:

$$(P_2 - P_1) = (\rho g)\,(y_2 - y_1). \tag{12.8}$$

This equation can also be written as

$$P_1 = P_2 - (\rho g)\,h, \tag{12.9}$$

where

$$h = y_2 - y_1.$$

Equation 12.9 can be useful for finding the pressure at any depth h in terms of the pressure on the surface P_{surf}. In this case, y_1 refers to the surface of the water; that is, $y_1 = 0$ and $y_2 = h$. Equation 12.9 then reduces to

$$P_2 = P_{surf} + \rho gh.$$

If no pressure other than atmospheric pressure is acting on the surface, then P_{surf} would equal atmospheric pressure. Changing the symbol from P_2 to P_h, in reference to its value at the depth h, the above equation becomes

$$P_h = P_{atm} + \rho gh. \tag{12.10}$$

In Equation 12.10, P_h is the total pressure, while the product ρgh represents the gauge pressure that was described in Example 12.2. This equation is in perfect agreement with the property stated in Section 12.2 (see Figure 12.1).

EXAMPLE 12.4

Find (a) the total pressure and (b) the gauge pressure at the bottom of a 5.50 m deep swimming pool.

Solution

a. Substituting for $h = 5.50$ m, $g = 9.80$ m/s², $\rho = 1.00 \times 10^3$ kg/m³, Equation 12.10 becomes

$$\begin{aligned} P_h &= 1.01 \times 10^5\ \text{N/m}^2 + (1.00 \times 10^3\ \text{kg/m}^3)\,(9.80\ \text{m/s}^2)(5.50\ \text{m}) \\ &= 1.01 \times 10^5\ \text{N/m}^2 + 0.539 \times 10^5\ \text{N/m}^2 \\ &= 1.55 \times 10^5\ \text{N/m}^2 \\ &= 1.53\ \text{atm}. \end{aligned}$$

b. The gauge pressure is

$$P_g = \rho gh$$
$$= (1.00 \times 10^3 \text{ kg/m}^3)\ (9.80 \text{ m/s}^2)\ (5.50 \text{ m})$$
$$= 0.539 \times 10^5 \text{ N/m}^2$$
$$= 0.534 \text{ atm.}$$

Comment

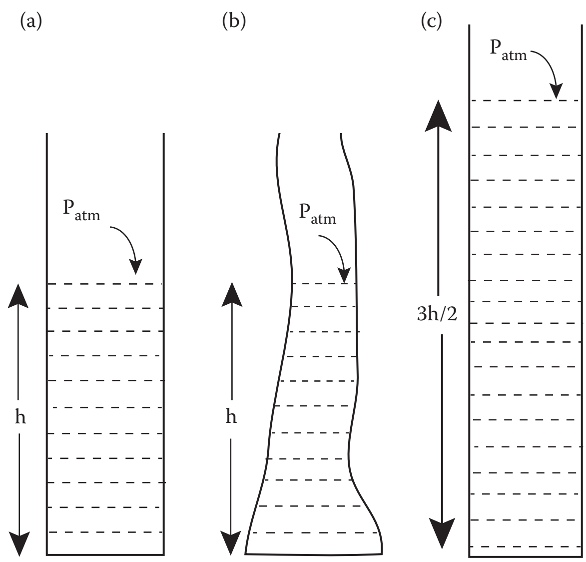
In stating the pressure P_h at higher altitudes, it is usual to associate that with the height h above earth's surface. The pressure P_h in terms of P_{atm} and h can be established by following an argument similar to that followed in deriving Equation 12.10. The relation in this case takes the form:

$$P_h = P_{atm} - \rho gh, \tag{12.11}$$

where ρ is the density of the earth's atmosphere, which if assumed constant, calculation of P_h becomes trivial.

EXAMPLE 12.5

Compare between the pressures on the bottom of the three containers shown in the figure below.



Solution

From Equation 12.10,

$$P_h = P_{surf} + \rho gh.$$

Aside from the atmosheric pressure on any open surface, the pressure at any point within a liquid depends only on its depth beneath the surface. Thus, for both, the tube in (a) and bottle-shaped container (b), the pressure is

$$P_{bottom} = P_{atm} + \rho gh,$$

while for tube (c), the pressure on its bottom is

$$P_{bottom} = P_{atm} + \rho g(3/2)h.$$

Analysis

The answers for tubes (a), (b), and (c) did not depend on shape nor geometry of the container. It is just the height of the liquid above the position of interest that determines the pressure.

12.5 PASCAL'S PRINCIPLE AND APPLICATIONS

Equation 12.10 shows that the pressure P_h at a depth h in a liquid can be increased if the pessure on the surface of the liquid is increased. It was Pascal (1623–1662) who first realized this property that can be stated: An external pressure applied to a static fluid is transmitted undiminished, throughout the fluid. This principle is used in numerous hydraulic applications. Among these are the hydraulic lifts and presses in car service workshops.

To explain the process of applying Pascal's principle, consider a simplified sketch of a hydraulic jack (Figure 12.3) that has a U shape with two sides, left and right of areas A and A′, respectively. Once a force F acts on the jack's left side, a pressure P is applied on the cylinder surface and on the liquid surface on that side. This is in addition to the atmospheric pressure acting on the fluid's surface on the same side. The pressure due to F is equal to

$$P = \frac{F}{A}, \tag{12.12}$$

from which

$$F = AP.$$

Since fluids transmit pressure to all of its parts undiminished, the pressure P transmits equally to the right side of the tube. Denoting that by P′, then

$$P' = \frac{F'}{A'}, \tag{12.13}$$

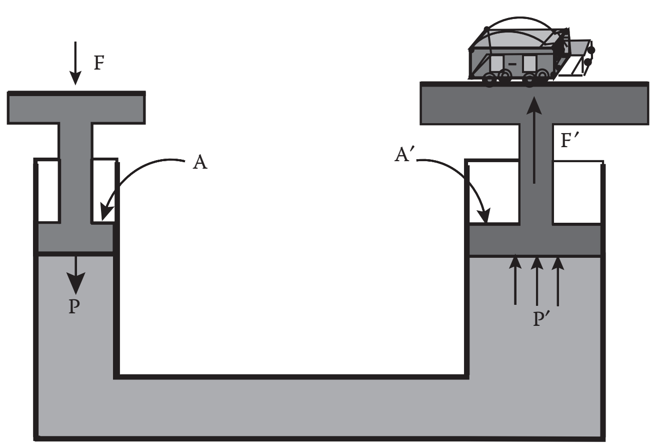


FIGURE 12.3 A schematic of a hydraulic jack that has a U shape. Its left and right sides top surface areas are A and A′, respectively.

from which

$$F' = A' P'.$$

As

$$P = P',$$

then Equations 12.12 and 12.13 give

$$\frac{F}{A} = \frac{F'}{A'}.$$

That is,

$$\frac{F'}{F} = \frac{A'}{A}.$$

This means that the ratio F′/F is determined by the ratio A′/A. From the above equation, F′ becomes,

$$F' = F\left(\frac{A'}{A}\right). \tag{12.14}$$

As the area A designates the side on which an input force, F, is applied, it is called the input area, A_{in}, while A′ is the area of the press where the heavier load is to be lifted, and is called the output area, A_{out}. The load represented by F′ is called the output force, F_{out}. Thus, Equation 12.14 may equivalently be reexpressed as

$$F_{out} = F_{in}\frac{A_{out}}{A_{in}}. \tag{12.15}$$

12.6 ARCHIMEDES' PRINCIPLE

Equation 12.10 indicates that the pressure of a liquid increases with depth, an experience evident to swimmers and divers. Thus for an object, in the form of a circular disk of mass m, flat side area A, and thickness t, immersed in a liquid of density ρ (Figure 12.4a), the force on its lower surface is larger than that on its upper side. This difference is due to an upward force called the buoyant force. In Figure 12.4a, the force on the upper side of the disk is

$$F_1 = P_1 A.$$

Using Equation 12.10, this becomes

$$F_1 = (P_{atm} + \rho g d_1)\,A.$$

Similarly, the force on disk's lower side is $F_2 = P_2A$, that is,

$$F_2 = (P_{atm} + \rho g d_2)\,A.$$

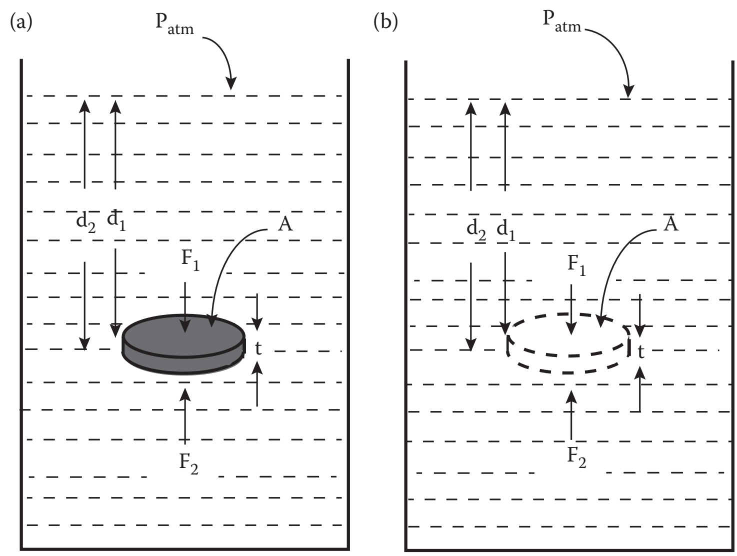


FIGURE 12.4 Analysis of the forces acting on the top and lower sides of (a) an object in the form of a circular disk of finite volume and mass immersed in a liquid; (b) a visualized volume of the liquid identical in form and volume to the disk and at the same depth.

The net force upward on the disk is equal to the buoyant force, F_b. Thus,

$$F_b = F_2 - F_1 = \rho g A\,(d_2 - d_1)$$

or

$$F_b = \rho g A t.$$

Since $At = V$ is the volume of the disk and and ρV is the mass m of a volume of the liquid equal to the volume of the disk (Figure 12.4b), then

$$F_b = \rho g A t = g\rho V$$

or

$$F_b = mg. \tag{12.16}$$

Thus,

$$F_b = W, \tag{12.17}$$

where W is the weight of the volume V of the liquid that is displaced by the disk. Thus, the buoyant force acting on the disk is equal to the weight of the displaced liquid, and it is that amount by which the apparent weight of the disk is diminished when immersed in the liquid.

EXAMPLE 12.6

Consider two containers filled with water where in one a piece of aluminum of volume 0.200 m^3 is hung from the free end of a spring scale and lowered until it is fullly submerged in the liquid, while a piece of lead of the same volume is similarly submerged in the other container. The densities of aluminum and lead are 2.70×10^3 and 11.35×10^3 kg/m^3, respectively (CRC, 70th Edition, 1989–1990, pp. B-68, B-23).

a. Find the buoyant force acting on each of the two pieces.
b. Determine the reading of the scale for each piece submerged in water?

Solution

a. As the aluminum piece is lowered in the first vessel, an equal volume of water will be displaced. The mass of the displaced water is

$$m_w = \rho V = (1.00 \times 10^3 \text{ kg/m}^3)\ (0.200 \text{ m}^3) = 0.200 \times 10^3 \text{ kg}.$$

The buoyant force on the aluminum is

$$F_b = W_w = m_w\ g = (0.200 \times 10^3 \text{ kg})\ (9.80 \text{ m/s}^2) = 1.96 \times 10^3 \text{ N}.$$

Since the volume of the lead piece is identical to the volume of the aluminum, the buoyant force on it is the same, that is,

$$F_b = W_w = m_w\ g = (0.200 \times 10^3 \text{ kg})\ (9.80 \text{ m/s}^2) = 1.96 \times 10^3 \text{ N}.$$

b. The masses and weights of the two pieces are:

Aluminum:

$$m = \rho_{al} V_{al} = (2.70 \times 10^3 \text{ kg/m}^3)\ (0.200 \text{ m}^3) = 5.40 \times 10^2 \text{ kg}.$$

So its weight in air,

$$W_{al} = mg = (5.40 \times 10^2 \text{ kg.})\ (9.80 \text{ m/s}^2) = 5.29 \times 10^3 \text{ N}.$$

Lead:

$$m = \rho_{lead} V_{lead} = (11.35 \times 10^3 \text{ kg/m}^3)\ (0.200 \text{ m}^3) = 2.27 \times 10^3 \text{ kg},$$

and its weight in air,

$$W_{lead} = mg = (2.27 \times 10^3 \text{ kg})\ (9.80 \text{ m/s}^2) = 2.22 \times 10^4 \text{ N}.$$

In water each piece will have its apparent weight diminished by an amount equal to the buoyant force acting on it. Thus, the reading of the scale for each will be:

Aluminum:

$$W_{scale} = W_{al\text{-}air} - F_b = 5.29 \times 10^3 \text{ N} - 1.96 \times 10^3 \text{ N} = 3.33 \times 10^3 \text{ N}.$$

Lead:

$$W_{scale} = W_{lead\text{-}air} - F_b = 2.22 \times 10^4 \text{ N} - 1.96 \times 10^3 \text{ N} = 2.02 \times 10^4 \text{ N}.$$

Analysis

The buoyant force exerted on an object by a fluid depends on its volume, not its mass or weight or shape. This asserts that it depends on volume of fluid displaced by the submerged object.

12.7 DYNAMICS OF FLUIDS: EQUATION OF CONTINUITY

The differences between properties of fluids and solids make fluid dynamics more involved than that of particles and rigid bodies. However, among fluids, liquids that are in steady state have properties that simplify its dynamics using mass and energy conservation principles.

For liquids such as water, where compressibility and viscosity are minimum, a streamline flow is assumed. This implies that water molecules running in a hose or a river pass any location with the same velocity as the molecules that passed that very location earlier. Thus for a flexible horizontal pipe of variable diameter (Figure 12.5), a stream of liquid in steady state, flowing from location (1) to (2), and passing through the cross-sectional area A_1 at a certain instance will pass through A_2 at a later time. Letting the average velocity of the liquid flow at locations (1) and (2) to be v_1 and v_2, respectively, the amount of mass Δm of liquid flowing in a small time interval Δt through A_1 is

$$\Delta m = \rho v_1 A_1 \Delta t,$$

where ρ is the density of the liquid. As this amount gets to location (2), incompressibility requires the mass flow rate to be constant; thus for an identical time interval Δt the same amount will pass through A_2, that is,

$$\Delta m = \rho v_2 A_2 \Delta t.$$

Equating the right hand sides in the above two equations gives

$$v_1 A_1 = v_2 A_2. \tag{12.18}$$

Notice, the SI units of vA are m^3/s, that is, volume over time. However, a more practical unit that is more common is liter per minute (L/min). That is why this quantity is called volume flux or volume flow rate, often denoted by Q. Equation 12.18 is known as the equation of continuity.

12.8 BERNOULLI'S EQUATION

In this section, a liquid of density ρ, incompressible with streamline flow and negligible viscosity, is presented using conservation of energy principle. Let the liquid flow in a pipe between locations (1) and (2), of heights, y_1 and y_2, respectively, measured from the ground (Figure 12.6). Again, let the speed of the liquid be v_1 and v_2 at (1) and (2), where the cross-sectional areas are A_1 and A_2, respectively.

Finally, consider a small element of length ΔL_1 at (1) and another ΔL_2 at (2) that extend over two volume segments ΔV_1 and ΔV_2. Each of the elemental lengths ΔL_1 and ΔL_2 is traversed by the flow in a time interval Δt such that

$$\Delta L_1 = v_1 \Delta t$$

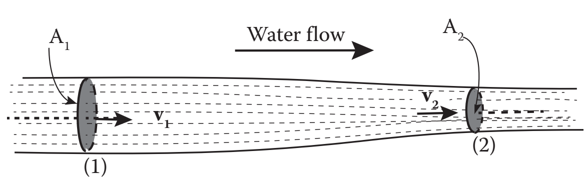


FIGURE 12.5 Demonstration of a horizontal stream of liquid flowing from location (1) to (2) with velocities v_1 and v_2, respectively. The cross-sectional areas are A_1 and A_2.

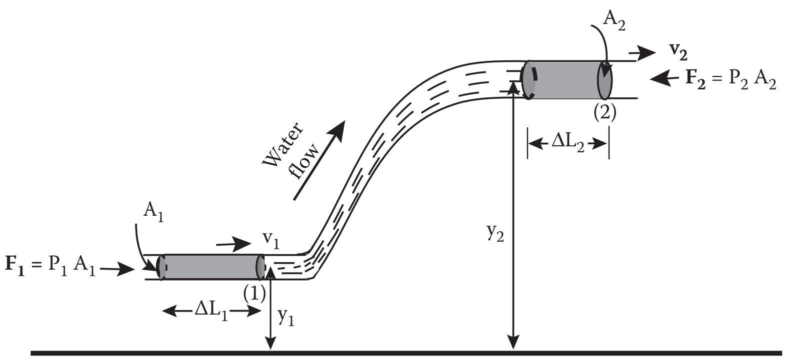


FIGURE 12.6 A selected tube of liquid flowing between two locations (1) and (2) with velocities v_1 and v_2, respectively. The tube at the two locations has areas A_1 and A_2, heights y_1 and y_2, and pressures P_1 and P_2, respectively.

and

$$\Delta L_2 = v_2 \, \Delta t.$$

Assuming the pressures on the two areas are P_1 and P_2, a force F_1 would be acting on A_1 and another F_2 on A_2. Notice that F_1 is due to the liquid on the left of the shaded segment at (1), and hence F_1 is parallel to the direction of the flow at (1), while F_2 is due to the inertia of the liquid on the right of the shaded segment at (2), and hence F_2 is opposite to the direction of the flow at (2). The values of the two forces are

$$F_1 = P_1 \, A_1$$

and

$$F_2 = P_2 \, A_2.$$

The work done on segment (1) is

$$\begin{aligned} W_1 &= F_1 \, \Delta L_1 \\ &= P_1 \, A_1 \, \Delta L_1. \end{aligned}$$

That is,

$$W_1 = P_1 \, \Delta V_1. \tag{12.19}$$

The work done on segment (2) is

$$\begin{aligned} W_2 &= -F_2 \, \Delta L_2 \\ &= -P_2 \, A_2 \, \Delta L_2. \end{aligned}$$

That is,

$$W_2 = -P_2 \, \Delta V_2. \tag{12.20}$$

The total work done on the moving segments ΔL_1 and ΔL_2 is

$$W = W_1 + W_2$$

$$= P_1 A_1 \Delta L_1 - P_2 A_2 \Delta L_2$$

$$= P_1 \Delta V_1 - P_2 \Delta V_2. \tag{12.21}$$

However, due to fluid incompressibility, the two volume elements ΔV_1 and ΔV_2 are equal. That is,

$$\Delta V_1 = \Delta V_2 = \Delta V, \text{ say}$$

$$W = (P_1 - P_{2)} \Delta V. \tag{12.22}$$

As the work done on a system by nonconservative forces like those resulting from the pressure of the liquid is equal to the change in its total mechanical energy, then,

$$W = \Delta E$$

$$= \Delta K + \Delta U.$$

ΔK and ΔU are the changes in the system's kinetic and potential energies, respectively. Thus, the above equation becomes

$$W = \frac{1}{2} \Delta m\,(v_2^2 - v_1^2) + \Delta mg\,(y_2 - y_1). \tag{12.23}$$

Substituting for the mass element, $\Delta m = \rho \Delta V$, and using Equation 12.22 for the left hand side of the above equation, it becomes

$$(P_1 - P_2)\Delta V = \frac{1}{2}\rho\Delta V(v_2^2 - v_1^2) + \rho\Delta Vg\,(y_2 - y_1)$$

or

$$(P_1 - P_2) = \frac{1}{2}\rho(v_2^2 - v_1^2) + \rho g\,(y_2 - y_1),$$

which upon rearrangement of terms becomes

$$P_1 + \frac{1}{2}\rho v_1^2 + \rho g y_1 = P_2 + \frac{1}{2}\rho v_2^2 + \rho g y_2, \tag{12.24}$$

known as Bernoulli's equation.

Comments:

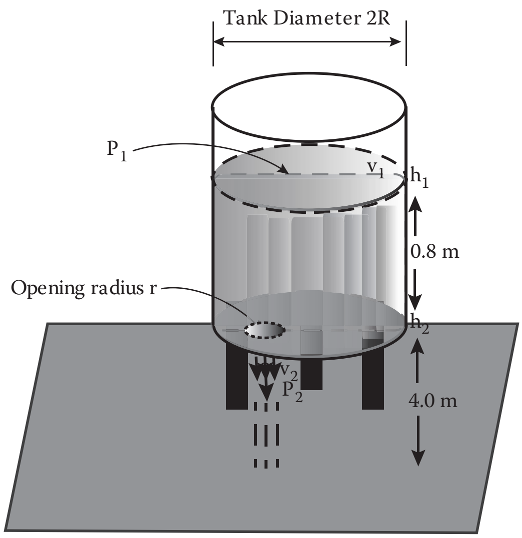
1. In the above treatment, the liquid was assumed ideal. In addition to ignoring viscosity, turbulence is also assumed absent. Thus, the liquid experiences no change in its internal energy. These are the conditions that make Equation 12.24 valid.

2. The presented argument leading to Equation 12.24 applies to any two positions 1 and 2 along the pipe. Thus, we may say that each of the sides in Equation 12.24 is constant. Thus, at any position in the streamline of the liquid,

$$P + \frac{1}{2}\rho v^2 + \rho g y = \text{constant.} \tag{12.25}$$

EXAMPLE 12.7

The figure below shows a cylindrical water reservoir 1.00 m in diameter. The water level is 0.80 m above the bottom. If a lid on a circular hole 2.0 cm in diameter in the bottom of the tank is removed, with what velocity will the water shoot out of the bottom of tank?



SOLUTION

Comparing the area of water coming out at the opening A_2 to the top water surface A_1 gives a ratio of

$$\frac{A_2}{A_1} = \frac{(\pi r^2)}{(\pi R^2)} = \frac{(0.010)^2}{(0.5)^2} = 4.0 \times 10^{-4},$$

which is very small. Thus, the speed of water molecules at the top surface can be considered stationary with respect to the bottom ones. Therefore, v_1 can be set to zero.

From Equation 12.24,

$$P_1 + \frac{1}{2}\rho v_1^2 + \rho g y_1 = P_2 + \frac{1}{2}\rho v_2^2 + \rho g y_2.$$

In the above equation, $P_1 = P_2 = P_{atm}$. Thus,

$$\frac{1}{2}\rho v_1^2 + \rho g y_1 = \frac{1}{2}\rho v_2^2 + \rho g y_2.$$

Substituting for the values, $y_2 = 0.00$, $v_1 = 0.00$, gives

$$(g)\,(y_1) = \frac{1}{2}\,(v_2^2) \qquad (12.26)$$

or

$$(9.80\ \text{m/s}^2)\,(0.80\ \text{m}) = \frac{1}{2}\,(v_2^2).$$

This gives

$$v_2 = 4.0\ \text{m/s}.$$

Analysis

1. Note that in Equation 12.26, v_2 can be expressed in terms of g and the height of water in the tank as follows

$$v_2 = \sqrt{2gy_1}.$$

   Recall that the above expression for v_2 is the same as that of a particle like object released from height h_1; v_2 is the object velocity just before it hits the ground. What a parallel!
2. A refreshing exercise can be practiced by working out this problem when the opening is placed on the side of an identical tank just above the bottom. If the tank is installed 4.0 m off the ground, at what horizontal distance away from the opening would the center of water shoot land? (see Problem 12.14).

PROBLEMS

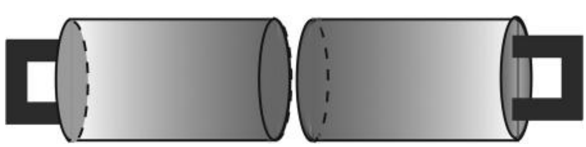
12.1 Consider a lady of 72.0 kg mass, wearing high heel shoes of 55 cm^2 total area of contact with the ground. Find the pressure her shoes exert on the ground.

12.2 A sky diver jumps from an airoplane 295 m above the ground.

   a. Find the difference in pressure between its values at the jumping altitude and value her body gets subjected to as she gets close to landing. (*Hint*: The density of air is 1.20 kg/m^3.)
   b. Knowing that the pressure on the earth's surface is 1.0 atm (1.01×10^5 Pa), find the pressure at the airoplane's altitude.

12.3 At what altitude would an airoplane experience a gauge pressure of 0.534 atm? Comment on the obtined answer by comparing it with the case described in Example 12.4.

12.4 Consider two identical hollow metallic cylinders, 28.0 cm in diameter. Each is open at one end and closed at the other (the figure below). The two, having thick walls, are brought together and partially evacuated to a pressure of 0.25 atm. Find the force needed to pull the two cylinders apart.



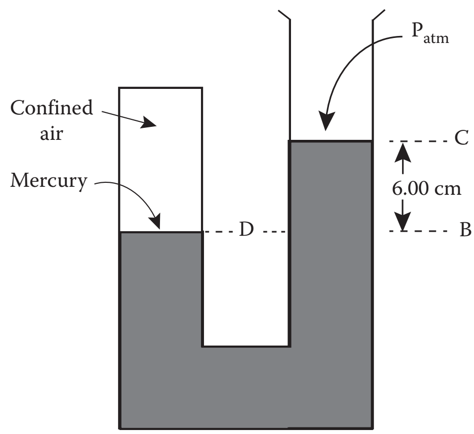
12.5 Consider cylindrical columns of (a) water and (b) air, 5.00 cm in diameter each. How high must each column be so as to enclose a masss of 1.00 kg would be?

12.6 Measurement of the volume of 100.0 g of: (a) water, (b) oil, and (c) alcohol is found to be 1.00×10^{-4} m^3, 1.05×10^{-4} m^3, and 1.24×10^{-4} m^3, respectively. Determine the density of the three liquids.

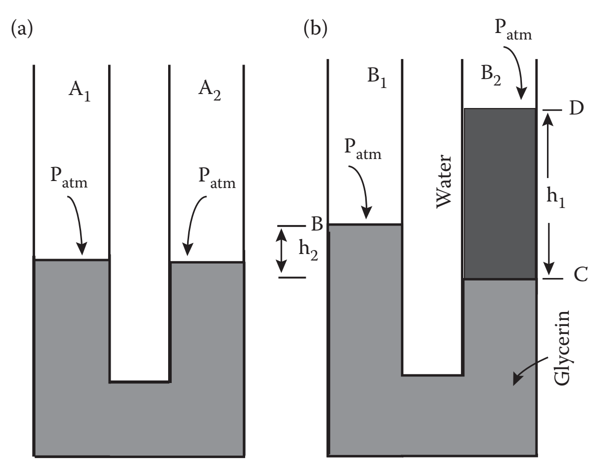
12.7 Two identical rectangular blocks of pine wood that measure 20.0 cm × 16.0 cm × 8.0 cm each are placed to float, one on water, and the other on oil. Determine how much of each will be submerged? (*Hint*: The density of wood is 0.500 kg/m^3, of water is 1.00×10^3 kg/m^3, and of oil is 0.875×10^3 kg/m^3.)

12.8 Consider an aluminum piece 0.400 m^3 hung from the free end of a spring scale. It is lowered into a water-filled vessel until it is fully submerged in the water. Knowing the density of aluminum is 2.70×10^3 kg, find the force needed to raise the aluminum piece from water.

12.9 In a lab investigation designed on the basis of Boyle's law, two glass tubes, one closed at one end and inverted, and another open at its two ends, are connected via a flexible horizontal rubber tube (not shown) (the figure below). If some mercury is poured into the open end of the tube, till the difference between the mercury levels B and C in the two tubes is 6.00 cm, determine the pressure on the mercury surface (level D) inside the inverted tube.

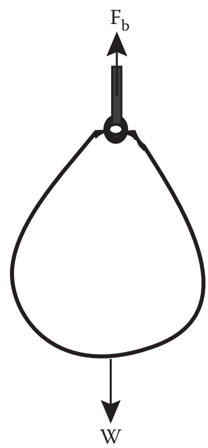


12.10 A U-shaped glass tube, open at both ends, is partially filled with glycerin (density 1.26×10^3 kg/m^3) (the figure below, part (a)). Water (density 1.00×10^3 kg/m^3) is poured in the right side to a height of h_1 = 17.0 cm above the glycerin level (part (b)). Find the height h_2 to which glycerin in the left side rises above its original level in the right side.



12.11 A balloon of 22.0 kg mass is filled with 52.0 m^3 helium of density $\rho = 0.179$ kg/m^3 (the figure below). Determine

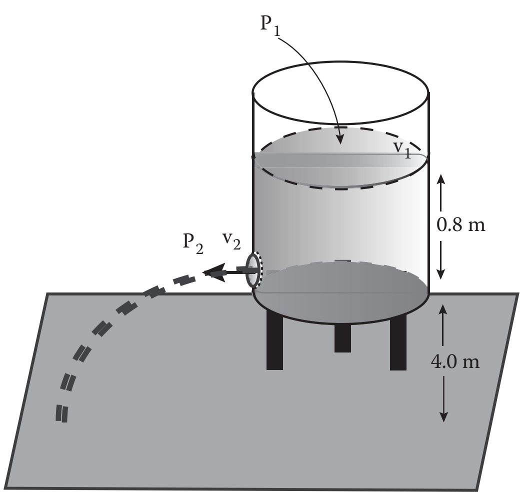
a. The buoyant force on the balloon as it starts to rise in the air
b. The force needed to hold it down during the final stage of preparation



12.12 Consider water flowing in a horizontal hose of a variable diameter, 2.54 cm at one location and 1.28 cm at another. Find the average speed in the wide part so that it flows out of the narrow part with a speed of 6.00 m/s. Consider the pressure within the hose equal to atmospheric pressure.

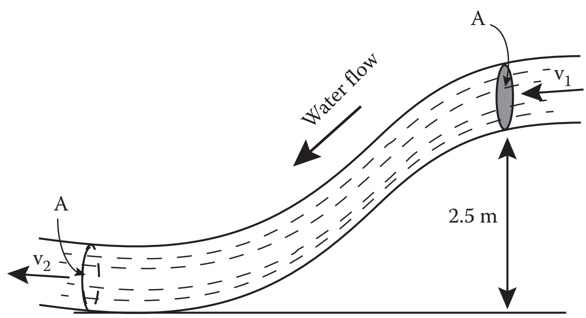
12.13 In a water cooling system, two horizontal pipes, one of 4.0 cm and another 8.0 cm diameter, are connected together such that water flows freely from the larger pipe to the narrower one. If water in the larger pipe is at 2.10×10^5 Pa pressure and 2.20 m/s speed, determine the (a) speed and (b) pressure of water in the narrow pipe.

12.14 Check Example 12.7 and work it out after a slight modification by having the opening (diameter 2.00 cm) on the side of the tank, slightly above its bottom (the figure below). If the tank, of diameter 1.00 m, is installed 4.00 m off the ground, how far away from the opening would the center of the stream of water land?



12.15 The figure below shows a segment of a uniform large hose with a 10.0 cm diameter. Water flows at a rate of 10.0 l/s. Knowing that the pressure at the lower side of the tube is 42 kPa, determine

a. The speed of water at both sides of the tube.
b. The pressure at the higher end of the segment.



13 Electric Forces and Fields

13.1 INTRODUCTION

Electric charge is one of the indefinables in physics along with displacement, mass, time, and thermodynamic temperature. An indefinable cannot be expressed in terms of a more fundamental quantity. Electric charge is inherent in the microscopic atomic structure of matter. The "planetary" model suggests and experiments corroborate that an atom consists of a relatively compact nucleus made up of two types of fundamental particles, protons, and neutrons. Orbiting the nucleus is a cloud of diffuse matter called electrons. Quantum theory predicts that these electrons cannot be localized, but it is often a mental convenience to think of them as classical point particles. Thus, point particle electrons orbiting a compact nucleus are analogous to a planetary system where planets orbit a sun.

It is possible to separate these particles by various means and a whole series of experiments have been conducted to determine their properties. Experiments reveal the following:

1. Electrons placed in close proximity measurably repel each other.
2. Protons placed in close proximity repel each other.
3. Electrons and protons attract each other.
4. There is no measurable attraction or repulsion of either electrons or protons by neutrons.
5. Neutrons and protons are much more massive than electrons, with $m_n = 1.675 \times 10^{-27}$ kg, $m_p = 1.673 \times 10^{-27}$ kg, and $m_e = 9.109 \times 10^{-31}$ kg.

It is reasonable to ask, "what intrinsic property do electrons and protons possess to make them attract or repel each other?" The intrinsic property is *electric charge*. Electrons and protons attract each other gravitationally also but that force is always attractive and is much weaker (10^{-39}) than electrostatic attraction or repulsion. Since protons repel other protons but attract electrons, the charge on a proton must be qualitatively different than that on an electron. To distinguish these two types, we define the charge on the proton to be *positive* and that on the electron to be *negative*. The neutron is electrically *neutral*; that is, it does not possess the intrinsic property of charge. The symbol e is used to designate the charge on a proton. It is found that the charge on the electron (−e) is equal in magnitude, but not in sign, to that on the proton (+e) and it is the smallest free charge found in nature. The SI unit of charge is the *coulomb* (C). By its definition

$$e = 1.60 \times 10^{-19}\ \text{C}.$$

Experiments indicate that charge is *quantized*, that is, all other charges found in nature are an integer multiple (N) of the charge found on the proton. Thus, if q represents any arbitrary charge, then $q = Ne$.

Atoms are electrically neutral; that is, they contain equal numbers of protons and electrons. If one or more of its electrons are removed, the atom has a deficiency of electrons and becomes positively charged. If a proton is removed from the atom's nucleus, generally much harder to achieve than removing outer electrons, the atom will also shed an outer electron and become a new neutral elemental atom. Atoms are microscopic. To acquire charges in large number, one needs to consider macroscopic matter.

13.2 CONDUCTORS AND INSULATORS: CHARGED OBJECTS

Experiments show that the innermost electrons in an atom are more strongly attracted to the positive protons in the nucleus than are the outermost valence electrons. Materials formed by atoms whose outermost electrons are only weakly bonded to the parent nucleus are called *electrical conductors*. The outer electrons are weakly shared by the nuclei, do not remain attached to the parent atom, and can drift through the conductor in the fashion of an electron gas. Most metals and ionic solutions are good conductors. Materials whereby the outer electrons are strongly attracted are called *electrical insulators*. Most ionic-bonded materials are good electrical insulators. Some examples are rubber, glass, and wood.

It is possible to charge, that is, transfer charge from one object to another, both insulators and conductors. As an example, suppose an originally neutral rubber (ebonite) rod is rubbed with an originally neutral piece of animal fur (Figure 13.1).

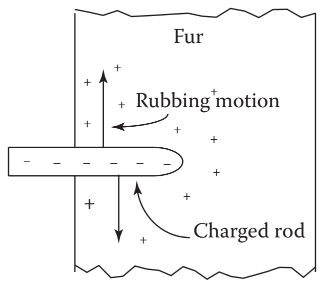


FIGURE 13.1 Originally neutral rubber (ebonite) and animal fur are rubbed together. The surface of the rubber acquires a net negative charge, the fur a positive charge.

The ebonite molecules have a higher electron affinity (electronegativity), that is, a higher tendency to form negative ions than does the fur. Thus, some electrons are transferred from the fur to the rod. The surface of the rod now has an excess of electrons and so is negatively charged, while the fur has a deficiency of electrons, that is more protons than electrons, which gives it a positive surface charge. This phenomenon shows itself when a rubber comb is used to comb dry hair. Thus, it is possible to charge insulators. The negatively charged rod can be used to charge a metal conducting sphere either positively or negatively.

When the rod is placed close, but not touching, the sphere, some valence electrons on the sphere are repelled resulting in a charge distribution as shown in Figure 13.2.

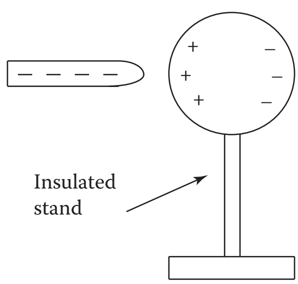


FIGURE 13.2 Net negative charges on the surface of an insulating rod repel some negative charges on the surface of a conducting sphere.

The charged rod has "induced" the skewed charge distribution on the sphere. Touching or rubbing the rod on the sphere allows some of the rod's surface charges to combine with the positive ions on the sphere thus neutralizing them (Figure 13.3).

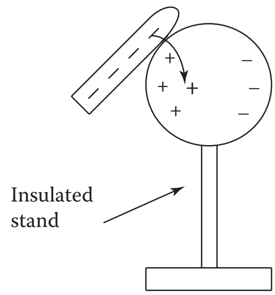


FIGURE 13.3 On contact, the rod transfers charge to the sphere.

If the rod is now removed, the excess negative charges on the sphere are redistributed by mutual repulsion (Figure 13.4). The sphere has been *charged by contact* and has the same sign as the charging rod. Alternately, bring the charged rod close to the sphere as in Figure 13.2. If one end of a conducting metal wire is attached to the sphere (Figure 13.5) while the other end is connected to the earth, some of the free electrons on the sphere will migrate along this "ground" wire and are distributed on the earth. Note that the earth, because of the concentration of metals on its surface, can be considered as a large source or sink for electrons, that is, many electrons can easily be taken from or dumped onto its surface. If the charged rod is removed and the ground wire disconnected, the sphere is left with a deficiency of electrons, that is, a net positive charge (Figure 13.6). The sphere has been *charged by induction* and has the opposite sign as the charging rod.

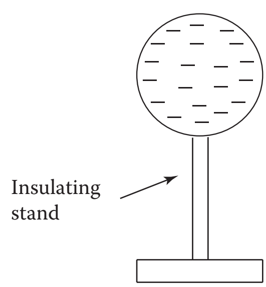


FIGURE 13.4 The net negative charges on the sphere redistribute themselves by mutual repulsion.

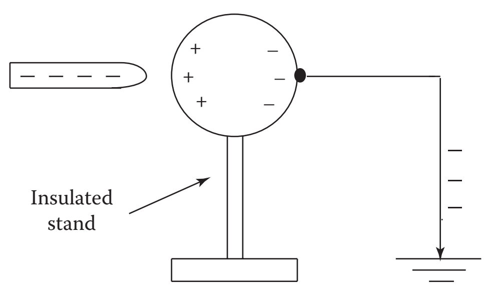


FIGURE 13.5 A setup for charging the sphere by induction.

An insulator cannot be charged by the induction method. Remember the ebonite rod was charged by rubbing *contact* with the fur. If the charged rod is held close to a sphere made of an insulating material, the rod may induce a "local" surface charge distribution of electric dipoles, that is, two

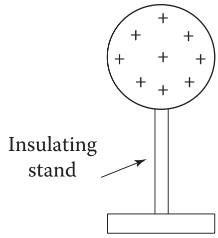


FIGURE 13.6 Removal of the grounding wire leaves a net positive charge on the conducting sphere.

equal and opposite charges separated by a small distance, and be attracted to them (Figure 13.7). This process explains the "static cling" of clothing due to the induced electric dipoles when clothes are tumbled in a dryer.

After the process of charging objects by contact or induction was discovered, fundamental experiments could be performed to reveal other properties of charges.

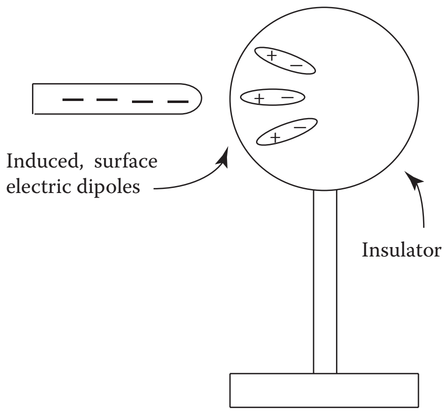


FIGURE 13.7 A charged rod may induce reorientation of electric dipoles on an insulating sphere.

13.3 COULOMB'S LAW

If an uncharged, metal sphere is placed on an insulating stand, it may be charged by induction or contact. If a second uncharged insulated identical metal sphere is touched to the first and then separated, the two spheres will equally share the original charge. Each will possess one-half the original charge. This process allows investigation of charges of known relative size.

Suppose two very small aluminum spheres of known mass, connected to insulating strings, both possess equal positive charges $q_1 = q_2 = N(+e)$. The strings are connected to circular rings that can not only slide along the horizontal rod but can also be locked in place (Figure 13.8). If the rings' separation is decreased, the angle θ increases due to the mutual repulsion of the "like" charges (Figure 13.9).

The charges, when at rest, are in static equilibrium and the forces acting on a charge are as shown in Figure 13.10. To determine the electric force **F** on either sphere, consider the horizontal and vertical force components.

$$\sum F_{\text{hor.}} = F - T\sin\theta = 0, \tag{13.1}$$

$$\sum F_{\text{vert.}} = T\cos\theta - W = 0. \tag{13.2}$$

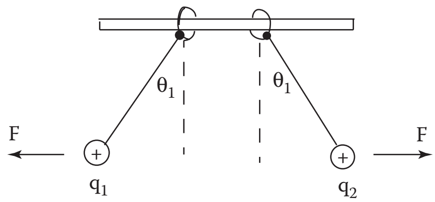


FIGURE 13.8 Charged spheres suspended on insulating strings repel each other.

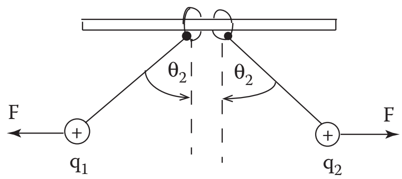


FIGURE 13.9 When the distance between the string supports is decreased, the angle with the vertical increases.

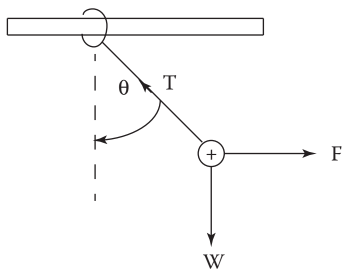


FIGURE 13.10 The free-body diagram of forces acting on a small-charged object.

Dividing Equation 13.1 by Equation 13.2 and rearranging yields

$$\frac{F}{W} = \frac{T\sin\theta}{T\cos\theta} = \tan\theta$$

or

$$F = W \tan\theta = mg \tan\theta. \tag{13.3}$$

So, the electric force on each charged sphere can be determined by measuring θ and the mass of the sphere. By varying the size of the charges, their sign and separation, the French physicist Charles Coulomb (1736–1806) through a series of experiments discovered the relationship that bears his name

$$F = \frac{kq_1q_2}{r^2}. \tag{13.4}$$

This is *Coulomb's law* (see Figure 13.11). The proportionality constant, in SI units, is $k = 8.99 \times 10^{-9}$ N m^2/C^2 in vacuum or air. It is conventional to express k in terms of ε_0 as $k = (1/4\pi\varepsilon_0)$ is called the *permittivity of free space* and has the numerical value of $\varepsilon_0 = 8.85 \times 10^{-12}$ C^2/Nm^2 in vacuum or air. Coulomb's law predicts that the electrostatic force is directly proportional to the product of the charges and inversely proportional to the square of their separation. So, F increases

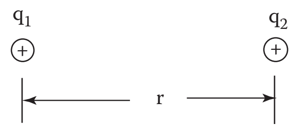


FIGURE 13.11 The geometry for illustrating the Coulomb force between two point charges.

if either q_1 or q_2 or both increase. As an example, doubling q_1 only doubles F, but doubling both q_1 and q_2 increases F by a factor of 4. The law also predicts that F decreases with increase in charge separation. As an example, for fixed q_1 and q_2, increasing the separation r to three times its original value results in a force that is only (1/9) as large as the original.

Equation 13.4 gives the magnitude of the vector force F. The vector form of Equation 13.4 is

$$\mathbf{F} = \frac{kq_1q_2}{r^2}\mathbf{r}_u. \tag{13.5}$$

Here, $\mathbf{r}_u$ is a unit vector directed along an imaginary line between the two charges. The signs of q_1 and q_2 determine the direction of the force.

For example, Figure 13.12a shows two positive charges, so the force on charge q_2 due to q_1 labeled $F_{2,1}$ is repulsive and directed away from q_1. Correspondingly, if q_2 is a negative charge, the force is attractive and directed toward q_1 (Figure 13.12b).

Equation 13.5 is a vector equation, so if more than two charges are interacting with each other the net force on any one of the charges can be found by the "addition of vectors" method.

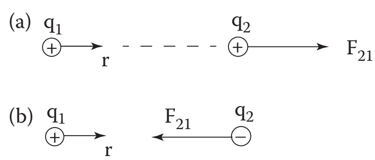
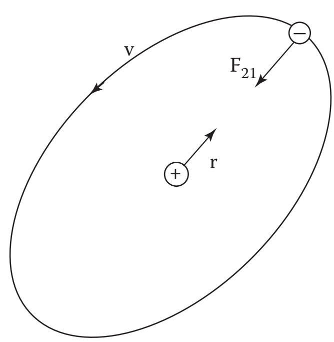


FIGURE 13.12 (a) The relative direction of a Coulomb repulsive force. (b) The relative direction of a Coulomb attractive force.

EXAMPLE 13.1: HYDROGEN ATOM (BOHR'S MODEL)

In Bohr's model, the electron of charge $q_2 = (-e)$ revolves about the nucleus of charge $q_1 = (+e)$ at a distance (first Bohr's radius) $r = 5.29 \times 10^{-11}$ m (the figure below). Determine the value of the Coulomb force that keeps the electron in orbit.



The planetary model of a hydrogen atom.

Solution

F_{21} is always directed radially inward toward the proton, even though it is continually changing its spatial direction. The magnitude is

$$F = \frac{kq_1q_2}{r^2} = \frac{(8.99 \times 10^9)(1.60 \times 19^{-19})(1.60 \times 10^{-19})}{(5.29 \times 10^{-11})^2}$$
$$= 8.22 \times 10^{-8} N = 1.85 \times 10^{-8} \text{ lbs.}$$

EXAMPLE 13.2

Compare the electrostatic force to the gravitational force between the electron and proton in the Bohr hydrogen atom.

Solution

The magnitude of the gravitational force is given by

$$F = \frac{GM_pM_e}{r^2}$$
$$= \frac{(6.67 \times 10^{-11})(1.67 \times 10^{-27})(9.11 \times 10^{-31})}{(5.29 \times 10^{-11})^2}$$
$$= 3.63 \times 10^{-47} N.$$

From the previous example, the magnitude of the electrostatic force is

$$F = 8.22 \times 10^{-8} \text{ N},$$

then

$$\frac{F(E)}{F(G)} = \frac{8.22 \times 10^{-8} N}{3.63 \times 10^{-47} N}$$
$$= 2.27 \times 10^{39}.$$

The electrostatic force is 10^{39} stronger, so gravitational forces are usually ignored in electrostatic calculations.

EXAMPLE 13.3

Given the charge configuration of the following figure (part (a)), determine the force F_{12} exerted on q_1 by charge q_2 and F_{13} by q_3.

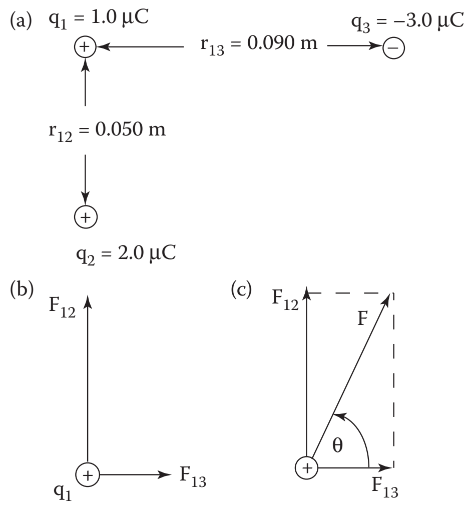
Solution

Charge q_2 is (+) so it exerts a repulsive force on q_1 of magnitude

$$F_{1,2} = \frac{kq_1q}{(r_{1,2})^2} = 7.2 \text{ N}.$$

Charge q_3 is (–) so it exerts an attractive force on q_1 of magnitude

$$F_{1,3} = \frac{kq_1q_3}{(r_{1,3})^2} = 3.3 \text{ N}.$$



(a) The charge locations for Example 13.3.
(b) Resolution of the forces into two components.
(c) The resultant force.

The magnitude of the net force on q_1 (the figure above, part (b)) is

$$F = \{[F_{1,2}^2] + [F_{1,3}^2]\}^{\frac{1}{2}} = 7.9 \text{ N}.$$

The direction of the net force (the figure above, part (c)) is

$$\theta = \tan^{-1}\left(\frac{F_{12}}{F_{13}}\right) = \tan^{-1}\left(\frac{7.2}{3.3}\right) = 65^{o}.$$

13.3.1 Other Considerations

Coulomb's law is an empirical law of wide generality:

1. It is easy to apply to point charges but more difficult to apply to larger irregular distributions. To see this, consider a single positively charged metal sphere of radius 10.0 cm supported by an insulating stand. If no other charges are in close proximity to the sphere, the charges on its surface will be uniformly distributed over its surface. If a second sphere, negatively charged, is brought close to the first, the mutual attraction of the charges will cause them to be redistributed on the surfaces so that neither sphere will have a uniform charge distribution (Figure 13.13).

   Even though the magnitudes if q_1 and q_2 are known, the choice of "r" in Coulomb's law becomes ambiguous. With certain geometries, the force could be calculated using calculus to integrate over the charge distributions. But Coulomb's law is easiest when applied to point charges or larger charge distributions if their separation is large compared with their size.
2. Coulomb's law is valid in vacuum, air, or other dielectric (nonconducting) materials such as wax, plastic, and glass. If point charges are imbedded in a dielectric (Figure 13.14), Equation 13.4 can be used to calculate the electric force between them provided the proper dielectric constant (ε) is used. Recall that k = $(1/4\pi\varepsilon_0)$. The new value of k is $(1/4\pi\varepsilon)$, where ε is the appropriate dielectric constant. As an example ε_0 (vacuum) = $\varepsilon_0 = 8.85 \times 10^{-12}$ C²/Nm² and ε(teflon) = $2.1\varepsilon_0 = 1.86 \times 10^{-11}$ $C^{2/Nm2}$.

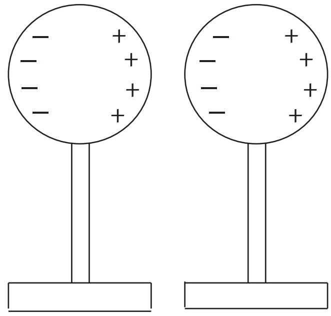


FIGURE 13.13 Redistribution of charges on conducting spheres.

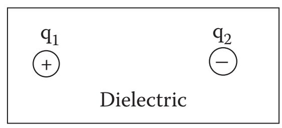


FIGURE 13.14 Point charges embedded in a dielectric medium.

13.4 ELECTRIC FIELD

Coulomb's law implies action-at-a-distance forces. It is reasonable to ask how does one proton "know" that another is in close proximity, that is, how is the repulsive force transmitted to the proton? Various mechanisms have been proposed, for example, the protons exchange virtual photons, but the process is not thoroughly understood. The action-at-a-distance can be explained by the concept of an *electric field*. A charged object can be thought of as altering or distorting the space surrounding it. This alteration is the electric field. If a second charged object is placed in this field, it experiences a force. The field can be quantified and mapped by use of an infinitesimally small positive test charge q_0. The test charge must be small enough so that it does not itself appreciably alter its own surrounding space thereby affecting the field that is to be mapped. The electric field **E**, due to charge distribution Q (Figure 13.15), is defined as

$$\mathbf{E} = \frac{\mathbf{F}}{q_0}. \tag{13.6}$$

Here, F is the force exerted on the test charge q_0 immersed in the field created by charge Q (Figure 13.15).

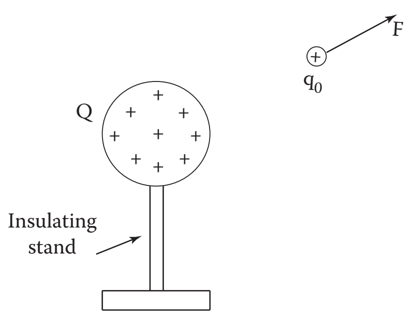


FIGURE 13.15 Charge Q produces an electric field **E**. Test charge q_0 responds to its effect.

Properties of E

1. **E** is a vector and points in the direction of the force **F** exerted on the positive test charge q_0.
2. The units of **E** are force per charge = Newtons per Coulomb = N/C.
3. The value of **E** at points outside the charge distribution Q can be "mapped" by placing q_0 at various locations and measuring the magnitude and direction of the electric force on it.
4. If **E** is known in a region of space, the force on an arbitrary charge in that region will be, from Equation 13.6,

$$\mathbf{F}_{(arb)} = \mathbf{E}q_{(arb)}. \tag{13.7}$$

5. The value of **E** at a distance r outside a small spherically symmetric charge distribution Q can be found from Coulomb's law. That is,

$$\mathbf{E} = \frac{F}{q_0} = \frac{\left[\frac{kQq_0}{r^2}\right]}{q_0} = \frac{kQ}{r^2}.$$

See Figure 13.16. Since the force is repulsive, **E** is directed radially outward from Q. Note that if Q was negative, **E** would be directed radially inward toward Q as would the force on q_0.

EXAMPLE 13.4

Determine the numerical value of the electric field at a position 10.0 cm from a point charge of value Q = +5.00 μC.

Solution

$$E = \frac{kQ}{r^2} = \frac{(8.99 \times 10^9\ \mathrm{Nm^2/C^2})(5.0 \times 10^{-6}\ \mathrm{C})}{(0.10\ \mathrm{m})^2 = 4.5 \times 10^6 (\mathrm{N/C})}.$$

EXAMPLE 13.5

Determine the value of the force on an arbitrary charge q = –3.0 μC at the distance r = 0.10 m from Q of Example 13.4.

Solution

$F = q_{arb}\,E = (-3.0 \times 10^{-6}\ \mathrm{C})(4.5 \times 10^6\ \mathrm{N/C}) = -14\ \mathrm{N}$. F is directed toward Q. Since **E** is a vector, its value at some point in space, due to several or more point charges, can be found by vector addition, that is, the principle of superposition.

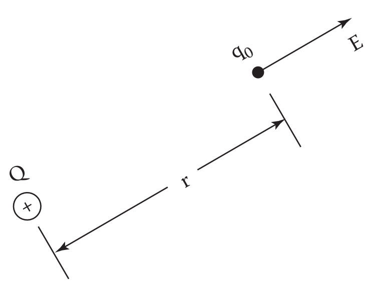
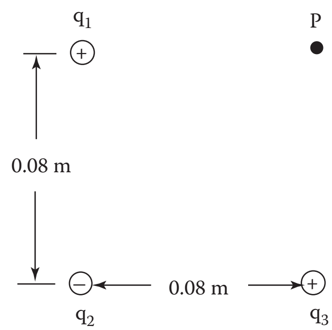


FIGURE 13.16 Depiction of some electric field properties.

EXAMPLE 13.6

Determine the value of **E** at the unoccupied corner point P of the square (the figure below), due to the charges $q_1 = +1.0\ \mu C$, $q_2 = -2.0\ \mu C$, and $q_3 = +3.0\ \mu C$.



Charge locations for Example 13.6.

Solution

$$E_1 = \frac{kq_1}{r_1^2} = \frac{(9.0\times 10^9)(1.0\times 10^{-6})}{(0.08)^2} = 1.4\times 10^6\ \text{N/C},$$

$$E_2 = \frac{kq_2}{r_2^2} = \frac{(9.0\times 10^9)(-2.0\times 10^{-6})}{(0.11)^2} = -1.5\times 10^6\ \text{N/C},$$

$$E_3 = \frac{kq_3}{r_3^2} = \frac{(9.0\times 10^9)(3.0\times 10^{-6})}{(0.08)^2} = +4.2\times 10^6\ \text{N/C}.$$

The net horizontal component is (the figure below, part (a))

$$E_x = E_1 - E_2 \cos 45°$$

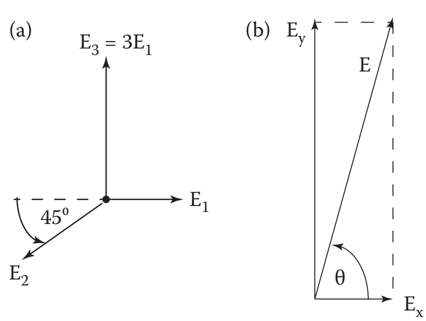
$$E_x = 1.4\times 10^6 - (1.5\times 10^6)(0.707)$$

$$= (1.4 - 1.1)\times 10^6 = 0.30\times 10^6\ \text{N/C}.$$

$$E_y = E_3 - E_2 \cos 45°$$

$$E_y = 4.2\times 10^6 - (1.5\times 10^6)(0.707)$$

$$= 3.1\times 10^6\ \text{N/C}.$$



(a) Resolution of the electric field into components for Example 13.6.
(b) The resultant electric field.

The magnitude of the net field is

$$E = [(E_x)^2 + (E_y)^2]^{1/2}$$

$$= [(0.3 \times 10^6)^2 + (3.1 \times 10^6)2]^{1/2}$$

$$= 3.1 \times 10^6 \text{ N/C (the figure above, part (b))}.$$

The net field direction is

$$\theta = \tan^{-1}\left(\frac{E_y}{E_x}\right) = \tan^{-1}\left(\frac{3.1 \times 10^6}{0.3 \times 10^6}\right) = \tan^{-1}(10)$$

$$\theta = 84°.$$

In discussing the definition of **E** (Equation 13.6), it was suggested that the spatial variation of **E** could be "mapped" by measuring the magnitude and direction of the force on q_0 at all points around Q. The great English physicist Michael Faraday (1791–1867) proposed a convenient way of visualizing **E** by means of *electric field lines*.

For a single charge (+Q), the electric field vectors **E** would be directed radially outward from (+Q) since that would be the direction of the force it exerts on ($+q_0$) (Figure 13.17a). For a charge (−Q), the **E** vectors would be directed radially inward toward (−Q) since the force on ($+q_0$) would be attractive (Figure 13.17b).

Two equal but opposite charges ($+q_1$) and ($-q_2$) create separate fields $\mathbf{E_1}$ and $\mathbf{E_2}$ at various spatial points (Figure 13.18). On a line joining the charges, both $\mathbf{E_1}$ and $\mathbf{E_2}$ at point C are directed toward ($-q_2$). The net **E** field at that point is the vector sum of $\mathbf{E_1}$ and $\mathbf{E_2}$. The force on q_0 at point C is then $\mathbf{F} = q_0\mathbf{E}_{net}$. At point D, $\mathbf{E_1}$ is directed radially outward from ($+q_1$) and $\mathbf{E_2}$ is directed radially inward toward ($-q_2$). The net field at point D would be the vector sum of $\mathbf{E_1}$ and $\mathbf{E_2}$. The force on q_0 at D would be in the same direction as $\mathbf{E}_{net}$. Indeed, the force on q_0 is always in the direction of $\mathbf{E}_{net}$, so the field lines are also called *electric lines of force*.

Some examples of field lines are shown. In Figure 13.19a and b, the lines provide a convenient visualization of the net force on a positive charge. From the mappings, it can be seen that the lines are closer to each other, that is, more lines per unit area, the force exerted on an arbitrary charge is greater. The force on a negative charge would, of course, be in a direction opposite to the direction of the field lines.

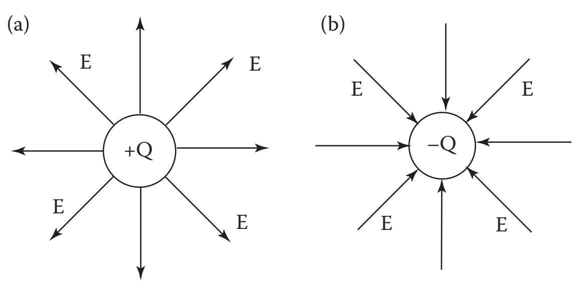


FIGURE 13.17 (a) Electric field lines emanate from positive charges. (b) Electric field lines converge onto negative charges.

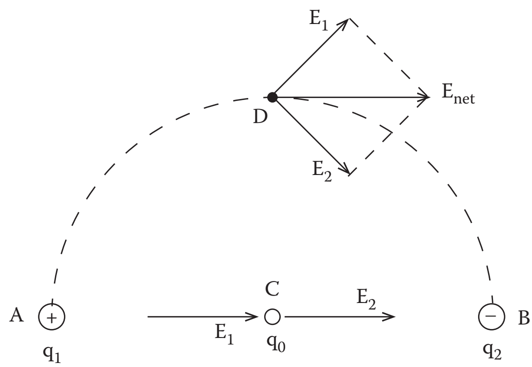


FIGURE 13.18 Net electric field directions are parallel to electric lines of force.

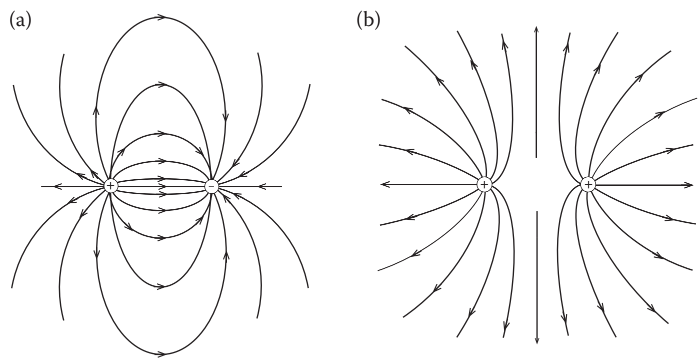


FIGURE 13.19 (a) Electric lines of force associated with two equal but opposite charges. (b) Electric lines of force for two equal like charges.

13.5 CONDUCTORS IN AN ELECTRIC FIELD

Suppose a metal conducting sphere is given an excess of electrons, that is, charged by contact as in Figure 13.4. How will these excess electrons distribute themselves in the sphere? Recall that they are nearly free and can move easily throughout the metal. These electrons repel each other by Coulomb's law and this repulsive force will drive them to the surface of the sphere where their separation from each other will be the maximum allowed and hence the force on each will be a minimum. Once on the surface, no additional movement of the electrons occurs. They are in static equilibrium. If the sphere instead has an excess of positive charge, that is, a deficiency of electrons, by similar argument this deficiency would reside in the surface atoms. This is equivalent to a surface distribution of positive charge. If the conductor is of a different geometry, for example, ellipse, long rod, or an irregular shape, Coulomb's law would also distribute the excess charge on their surface. So, in general, if charge is not moving through the conductor, that is, it is in electrostatic equilibrium, *any excess charge resides on the conductor's surface*.

Inside a conductor with or without an excess of surface charge, the electric field must be zero. If the field were not zero, then it would exert a force on the nearly free conduction electrons causing them to move through the conductor. The electrons would then not be in equilibrium. Thus, in equilibrium, *the electric field at any point interior to a conductor is zero*. So excess charge on the surface does not create an electric field inside a conductor. Also, if the conductor is hollow, that is, a spherical shell, E inside would be zero. The shell's interior would be "shielded" from external charges. This shielding property is exploited in sensitive electronic circuits by covering some of the circuit elements with a thin metal box, usually made of aluminum.

Lastly, what is the value of the electric field inside a conductor that is immersed in an external electric field? Consider two thin large conducting plates that have been uniformly and oppositely charged. Electric field lines emanate from positive and terminate on negative charges. If the plates are relatively large compared with their separation, the field lines in the space between them are approximately parallel (Figure 13.20). If a conducting sphere is placed in this field, the field will induce the surface charge distribution shown in Figure 13.21. Under electrostatic conditions, the electric field inside the sphere must be zero. Additionally, the field lines that emanate and terminate on the induced surface charges must be perpendicular to the conductor's surface. If this were not so, these vector field lines would have a component parallel to the conducting surface. This would lead to charge flow on the sphere's surface. This shielding property is exploited in sensitive electronic circuits. If an electronic component is affected adversely by external electric fields, it is placed in a metal box called a "Faraday cage." Similarly, the central conducting wire used in a television antenna connection is constructed with a braided metal sheath to shield the incoming signal. Such wires are called "coaxial cables."

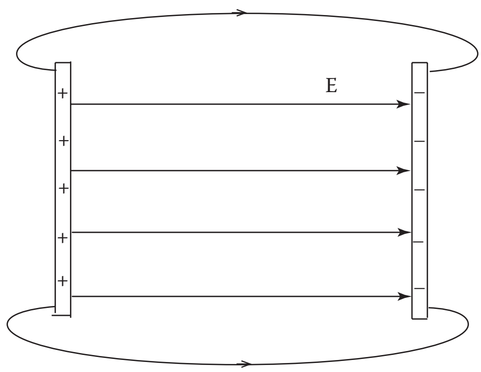


FIGURE 13.20 The approximately spatially uniform electric field between two charged, closely spaced parallel conductors.

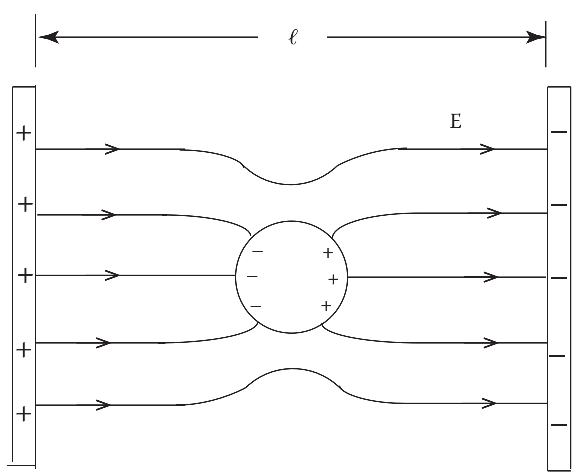


FIGURE 13.21 A conducting sphere immersed in an originally uniform electric field.

13.6 GAUSS' LAW

The electric field surrounding a point charge q can be found from Coulomb's law and was shown to be $E = kq/r^2$. The electric field due to many charges arranged together in a more complicated charge distribution, for example, a line of charges on the surface of a long nonconducting rod or charges uniformly dispersed throughout an insulating sphere, cannot be calculated via the point charge equation. Two common methods to determine E due to such charge distributions are as follows:

1. Brute force use of the point charge equation
2. Gauss' law

In the brute force method, the total charge is written as a sum of infinitesimal charges Δq. The Δqs are expressed in terms of the geometry of the distribution, each with its appropriate distance r to the field point. As an example, for a line of charge on a rod with uniform charge per unit length λ, the increment Δq would be $\Delta q = \lambda \Delta l$, where Δl is an increment of length. The contribution to the field would be $\Delta E = k\Delta q/r^2 = k\lambda\Delta l/r^2$. The Δls would be summed, using calculus, to yield the electric field due to the line of charge.

The second method, Gauss' law, is an elegant method for calculating electric fields and gives additional insight into the nature of electrostatics. It also shows some of the analogies between electric and gravitational fields.

To develop Gauss' law, recall that electric field lines or lines of force are associated with point charges and charge distributions. Recall also that in the regions where these lines are closer together, that is, more field lines per unit area, the field is stronger than in regions where the field lines are less dense. The field lines are then associated with a quantity Φ called the *electric flux*. The field E in any region of space can then be associated with this electric flux by the definition:

$$E = \frac{\Delta\Phi}{\Delta A_\perp} = \frac{F}{q_0}. \tag{13.8}$$

Here, $\Delta\Phi$ is the flux (number of lines of force) passing through an infinitesimal element of area ΔA that is perpendicular to the lines of force. To assure that the incremental area is perpendicular to **E**, that is, the projection of (ΔA) along **E**, Equation 13.8 is written as

$$\Delta\Phi = E(\Delta A)\cos\alpha. \tag{13.9}$$

Here, α is the angle between (ΔA), which is really a vector that is perpendicular to the physical area by definition and the direction of **E** (see Figure 13.22).

The total flux passing through an extended surface is acquired by adding the increments of flux, that is,

$$\Phi = \sum_{\text{surf}} \Delta\Phi = \sum_i E(\Delta A)_i \cos\alpha_i. \tag{13.10}$$

Now, consider the simplest case of the total flux associated with a point charge q. The surface to be used in Equation 13.10 is not a real physical surface. It is arbitrary and is indeed a mental construct called a *Gaussian surface*. Choose the surface to be a sphere of radius r with the charge q at its center (Figure 13.23). Notice that since **E** is directed radially outward from q, it is perpendicular

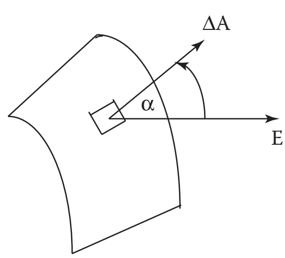


FIGURE 13.22 The incremental area vector (ΔA) is perpendicular to the surface of the physical area.

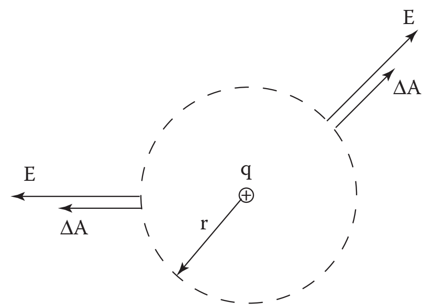


FIGURE 13.23 The imaginary spherical Gaussian surface surrounds and is centered at q.

to the Gaussian surface and has the same numerical value

$$E = \frac{kq}{r^2}$$

at all points on the surface. Note that ΔA is also perpendicular to the surface at all points on it and is thus parallel to the direction of **E** at all points. So the angles α_i are zero at all points on the surface. These conditions make the sum in Equation 13.10 easy to evaluate. Thus, since **E** and α_i are the same at all points on the Gaussian surface, they can be factored out of the sum. The remaining sum of all the incremental areas $(\Delta A)_i$ yields the surface area of the Gaussian sphere. Equation 13.10 becomes

$$\Phi = (\cos 0^\circ)\sum_i E(\Delta A)_i = \frac{kq}{r^2}(4\pi r^2), \tag{13.11}$$

$$\Phi = (4\pi\ kq) = 4\pi\left(\frac{1}{4\pi\varepsilon_0}\right)q, \tag{13.12}$$

$$\Phi = \sum_{\text{surf}} \Delta\Phi = \sum_i E(\Delta A)_i \cos\alpha_i = \frac{q}{\varepsilon_0}, \tag{13.13}$$

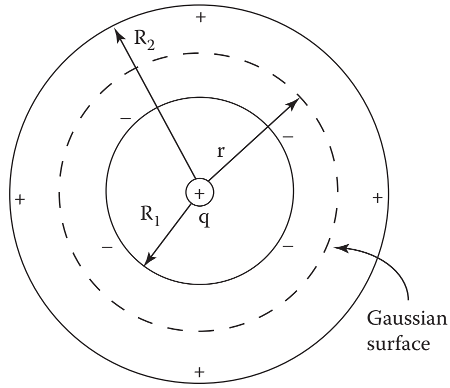
So, the total flux through the Gaussian surface is (q/ε_0), that is, the charge enclosed divided by ε_0. Equation 13.13 is *Gauss' law.*

Although a simple case, namely a point charge and a concentric spherical Gaussian surface, was used to establish Gauss' law, it can be shown to hold for arbitrary charge distributions and Gaussian surfaces. The Gaussian surface must be a *closed* surface and q is the *net* charge enclosed in it.

EXAMPLE 13.7

A hollow conducting sphere contains a positive point charge at its center (the following figure). Use Gauss' law to determine the electric field

a. Inside the hollow, $r < R_1$
b. Inside the conducting shell, $R_1 < r < R_2$
c. Outside the sphere, $r > R_2$



The Gaussian surface for Example 13.7.

Solution

a. Choose the mentally constructed Gaussian surface to be a sphere of radius $r < R_1$ centered on q. The flux through this closed surface is

$$\Phi = \sum_i E(\Delta A)_i \cos \alpha_i = \frac{q_{enclosed}}{\varepsilon_0}.$$

As mentioned previously, **E** and $(\Delta A)_i$ are perpendicular to the surface and parallel to each other at all points on the Gaussian surface, so $\alpha_i = 0$, and

$$\Phi = EA = \left(\frac{q}{\varepsilon_0}\right) = E(4\pi r^2)$$

or

$$E = \frac{q}{4\pi\varepsilon_0(r^2)}, \quad r < R_1.$$

b. Choose the Gaussian surface to be a sphere centered on q. Note that q at the center induces an equal magnitude negative charge on the inner surface of the conductor and a concomitant equal positive charge on the outer surface. The net charge within the Gaussian surface is

$$q_{net} = q_{center} + (-q) = 0, \quad R_1 < r < R_2.$$

So,

$$\Phi = \sum_i E(\Delta A)_i \cos \alpha_i = \frac{q_{enclosed}}{\varepsilon_0} = 0$$

and $E = 0$ inside the conductor, as it should under static conditions.

c. Choose the Gaussian surface to be a sphere centered on q with $r > R_2$. The net charge within the Gaussian surface is

$$q_{net} = q_{center} + (-q) + (+q) = +q, \quad r > R_2$$

and

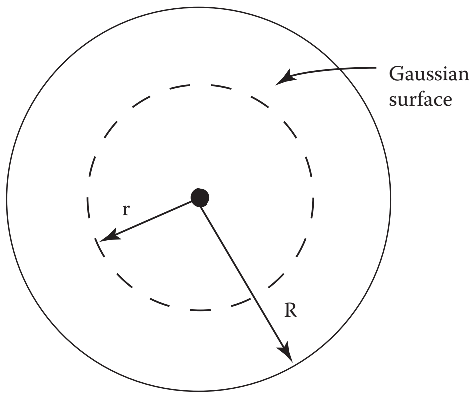
$$\Phi = \sum_i E(\Delta A)_i \cos \alpha_i = \frac{q_{enclosed}}{\varepsilon_0}.$$

By the same analysis used in part (a), the electric field is

$$E = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}, \quad r > R_2.$$

EXAMPLE 13.8

A solid insulating sphere is constructed with a uniform charge distribution ρ C/m³ throughout its volume (the figure below). Determine the electric field both (a) inside and (b) outside the sphere.



The Gaussian surface inside an insulator with a spatially uniform charge distribution.

Solution

a. Inside the sphere, choose the Gaussian surface to be a concentric sphere with $r < R$. The charge enclosed is

$$q_{enclosed} = \frac{\text{charge}}{\text{volume}}(\text{volume enclosed}) = \rho\frac{4\pi r^3}{3}.$$

Then,

$$\Phi = \sum_i E(\Delta A)_i \cos \alpha_i = \rho\frac{4\pi r^3}{3} = E(4\pi r^2)$$

or

$$E = \frac{\rho r}{3\varepsilon_0}, \quad r < R. \tag{13.14}$$

Note that E increases linearly as r increases.

b. Outside the sphere, choose the Gaussian surface to be a concentric sphere with r > R. The charge enclosed is the total charge possessed by the insulating sphere. Then,

$$\Phi = \sum_i E(\Delta A)_i \cos\alpha_i = \frac{q_{enclosed}}{\varepsilon_0} = \rho\frac{4\pi R^3}{3\varepsilon_0},$$

$$EA_{g.s.} = E(4\pi r^2) = \rho\frac{4\pi R^3}{3\varepsilon_0},$$

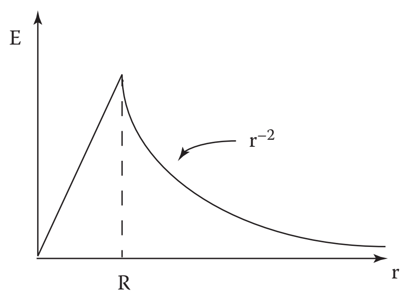
or

$$E = \frac{\rho R^3}{3\varepsilon_0}\frac{1}{r^2}. \tag{13.15}$$

Outside the sphere, **E** decreases as r increases and falls off in the same fashion as for a point charge. A plot of **E** inside and outside the sphere is shown in the figure below.

Note that E has its largest value when r = R, that is, on the surface of the sphere. When r = R, Equations 13.14 and 13.15 yield the same value

$$E = \frac{\rho}{3\varepsilon_0}R.$$



For the charged sphere of Example 13.8, the electric field, inside, increases linearly with increasing distance. Outside the sphere, E decreases as the reciprocal of the square of distance.

It is instructive to use Gauss' law to compare electrostatic to gravitational fields. The action-at-a-distance nature of gravitational forces can be explained in terms of a gravitational field g defined as

$$g = \frac{F_{grav}}{m_0} \tag{13.16}$$

gravitational force on a small test mass, that is, for a test mass above the earth. Equation 13.16 gives

$$g = \frac{(GM_e m_0/r^2)}{m_0} = \frac{GM_e}{r^2}. \tag{13.17}$$

A gravitational flux Φ_G can be defined in the same fashion as the electric flux Φ_E of Equation 13.13. Analysis similar to Example 13.2 would yield

$$g = \frac{\rho}{3}r \quad r < R. \tag{13.18}$$

Here, ρ is isotropic mass per volume. Also,

$$g = \frac{\rho R^3}{3}\frac{1}{r^2} \quad r > R. \tag{13.19}$$

These equations predict that g and hence weight, W = mg, are greatest on the earth's surface and is equal to zero at its center.

Gauss' law is a beautiful and powerful tool for determining the electric fields.

PROBLEMS

13.1 A metal sphere has an initial charge density of +6.00 μC placed on it. Determine the net charge on the sphere after an additional $3.00 \times 10^{+13}$ electrons have been placed on it.

13.2 Three stationary charges are affixed to a rectangular coordinate system. They are $q_1 = +16.0$ μC on the y axis at y = 2.50 cm, $q_2 = -9.00$ μC at the origin, and $q_3 = +39.0$ μC on the x axis at x = 3.50 cm. Determine the magnitude and direction of the electrostatic force on charge q_3.

13.3 Determine the magnitude of the force on a point particle of charge 3.00×10^{-6} C immersed in a uniform electric field of magnitude 2.50×10^5 N/C.

13.4 Determine the magnitude of the electric field at a distance of 1.50 m from a 4.00×10^{-6} C point charge.

13.5 Two charged objects attract each other with a force F. If the charge on one object is doubled and the separation between them is also doubled, the force between them is

a. 16F
b. 4F
c. F
d. F/2
e. F/4

Establish your answer.

13.6 An element of area with magnitude $\Delta A = 1.50 \times 10^{-4}$ m² is directed 25.0° with respect to the uniform electric field passing through it. The electric field magnitude is $E = 6.75 \times 10^{-4}$ N/C. Determine the element of electric flux (ΔΦ) passing through the area element.

13.7 Two spherical conducting shells have a common center. The radius of the inner shell is $r_1 = 7.00$ cm and the outer shell is $r_2 = 14.0$ cm. A charge of $q_1 = -1.50$ μC is distributed uniformly over the inner sphere and $q_2 = +6.00$ μC is spread uniformly over the outer shell. Determine the magnitude and direction of the electric field at a distance of (a) 5.00 cm, (b) 10.0 cm, and (c) 20.0 cm from the common center.

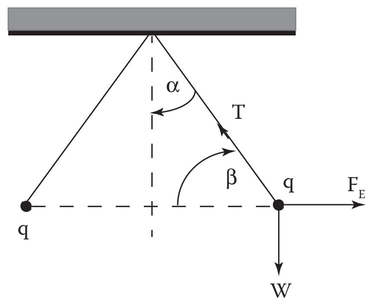
13.8 Two metal conducting spheres, each of radius r = 8.00 cm, have charges uniformly distributed over their surfaces. The charge on sphere one is $q_1 = +7.00$ μC and on sphere two is $q_2 = -2.60$ μC. The spheres are brought into contact and then separated. Determine the charge per area on each sphere after separation.

13.9 Two equally charged objects, separated by a distance of 50.0 cm, are found to exert an electrostatic force of 1.25 N on each other. Determine the amount of charge on each object.

13.10 The electric field at a certain point in space is found to be 1500 N/C in the (+x) direction. Determine the magnitude and direction of the acceleration of an electron released at that point.

13.11 A rectangular flat plate (0.120 m × 0.390 m) is placed in a uniform electric field of 570 N/C. (a) At what orientation of the plate will the electric flux through it be a maximum and a minimum? (b) Determine the numerical values of the maximum and minimum flux.

13.12 Two small spherical insulators, each of mass $m = 7.00 \times 10^{-2}$ kg and equal surface charge, are hung by light strings and repel each other such that each string is at an angle β with the horizontal. The forces acting on each charge are the Coulomb force (F_E), the gravitational force (W), and the tension (T) in the string. If $\beta = 30.0°$, determine the value of (F_E).



13.13 The two protons in the nucleus of the helium atom have a separation of approximately 3.00×10^{-15} m. Determine (a) the magnitude of the electrostatic force that each proton exerts on the other, (b) the magnitude of the mutual gravitational force between the protons, and (c) the ratio of these two forces.

13.14 Two charges, $q_1 = +12.0$ μC and $q_2 = -3.00$ μC, are fixed in space with a separation of 2.00 m. Determine the point, along a line through the charges, and relative to q_1, where the net electric field is zero. (*Hint:* q_2 is negative, so $\mathbf{E}_2$ is directed toward q_2.)

13.15 If the earth had a net charge of (+Q) and the moon had a net charge of (−Q), what should be the value of Q such that the electrostatic force between earth and moon is equal in magnitude to the average gravitational force between them?

14 Electric Potential Energy and Potential

14.1 INTRODUCTION

The previous chapter dealt with the phenomenon and properties of electric charge. The charges were generally at rest, that is, static, so the study was of electrostatics. We now address the motion of electric charges immersed in an electric field.

14.2 POTENTIAL ENERGY

Consider a large thin insulating plate uniformly charged with positive charges and a similar plate charged negatively. If these plates are separated by a distance ℓ that is small compared with the dimensions of the plates, then an approximately uniform electric field will exist between the plates (Figure 14.1). Suppose a positive charge (+q) is placed at point B a distance d_2 from the positive plate and momentarily held there by some external agent. If that agent now quasistatically moves q, initially at rest, against the E field to point A at position d_1, it will work on the charge and thus changes its potential energy.

The work done by the agent is

$$W_{BA} = \sum F \cdot \Delta x = F(d_1 - d_2)\cos\theta. \tag{14.1}$$

Here, θ is the angle between the applied force **F** and the displacement Δx and θ = 0 since **F** and **Δx** are parallel. Since q was moved quasistatically, no work was invested in changing its kinetic energy.

Correspondingly the work done by **E** on q while it was moved quasistatically from d_2 to d_1 is

$$W_{BA} = \sum F_E \cdot \Delta x = |F_E||(d_1 - d_2)|\cos\theta. \tag{14.2}$$

Here, θ = 180° since $\mathbf{F_E}$ and **Δx** are antiparallel, so $(W_E)_{BA}$ is negative that is,

$$W_{BA} = \sum F_E \cdot \Delta x = -|Eq|\,|(d_1 - d_2)|. \tag{14.3}$$

Now reverse the path direction. The work done by **E** when q is moved quasistatically from d_1 to d_2, with assistance by the external agent is

$$W_{AB} = \sum F_E \cdot \Delta x = |Eq||d_2 - d_1|\cos\theta. \tag{14.4}$$

Here, θ = 0° since $\mathbf{F_E}$ and **Δx** are parallel. So,

$$W_{AB} = +|Eq||d_2 - d_1|. \tag{14.5}$$

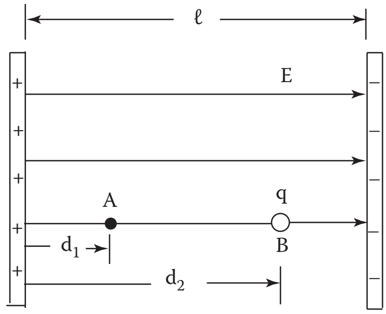


FIGURE 14.1 The work done on a point charge by a uniform electric field.

The work done by **E** when q is moved around the closed path d_2 to d_1–d_2 is

$$W_{BA} + W_{AB} = -|qE||d_1 - d_2| + |qE||d_2 - d_1| = 0. \tag{14.6}$$

Thus, it is seen that the work done by an electric field on a charged object is *conservative*, that is, its value is zero when computed around a closed path. Further, the work is independent of the path traversed by q. If q is started at d_2, moved by any path to d_1 and then back to d_2, the work done by **E** on q will be zero. This property is another similarity between electric and gravitational fields.

In Chapter 6, it was stated that the work done by a conservative force can be expressed as a negative change in potential energy. Here, the PE is *electric* potential energy. Equation 14.3 can be written as

$$W_{BA} = -|Eq||d_1 - d_2| = -\Delta PE. \tag{14.7}$$

Since both sides of Equation 14.7 are negative, ΔPE is positive. But, $\Delta PE = (PE)_f - (PE)_i$, that is, the final minus the initial values of PE. So when q was moved from d_2 to d_1, its PE increased, that is, its PE at d_1 is greater than at d_2. Similarly, Equation 14.5 can be written as

$$W_{AB} = +|Eq||d_2 - d_1| = -\Delta PE. \tag{14.8}$$

For this motion, ΔPE is negative, so when q was moved from d_1 to d_2, its PE decreased. Recognize that when q was moved against the direction of **E**, its PE was increased, and when it was moved in the direction of **E**, its PE was decreased. These processes were quasistatic and required the action of an external agent to keep the charge from acquiring a kinetic energy. They are analogous to the gravitational case of lifting an object off the floor and placing it on a shelf to increase its gravitational PE. The reverse would be to lift the object from the shelf and slowly set it back on the floor to decrease its gravitational PE. In both cases, ΔKE = 0.

In Figure 14.1, the charge if released from point A would accelerate toward point B. So, its PE would decrease and its KE would increase, similar to a falling object.

EXAMPLE 14.1

Suppose the charged object in Figure 14.1 has a mass of 4.18×10^{-8} kg, a charge of 4.00 μC, and E = 80.0 N/C. Determine the velocity of the object at a time t = 2.00 s after it is released from rest at point A.

Solution

The time t and v_0 are given, so the equation $v_f = v_0 + at$ could be used if the acceleration were known. The acceleration can be found from

$$F = ma = qE$$

or

$$a = \frac{Eq}{m} = \frac{(80.0\ \text{N/C})(4.00 \times 10^{-6}\ \text{C})}{4.18 \times 10^{-8}\ \text{kg}} = 7.66 \times 10^{3}\ \text{m/s}^2.$$

So,

$$v_f = 0.00 + (7.76 \times 10^3\ \text{m/s}^2)(2.00\ \text{s}) = 1.53 \times 10^4\ \text{m/s}.$$

14.3 ELECTRIC POTENTIAL

Equations 14.7 and 14.8 show that the work done by E on q and hence the change in electric PE depend on the magnitude and sign of q. It is useful to characterize the energy at locations in an **E** field in terms of a charge-specific quantity called the *electric potential*. It is defined as the electric PE per unit charge, that is,

$$V \equiv \frac{PE}{q}. \tag{14.9}$$

Since PE is expressed in joules and charge in coulombs, the SI unit is joules per coulomb or volts. Equation 14.9 implies that a change in electric PE, that is, ΔPE, can be expressed in terms of a change in potential ΔV or a voltage difference.

Referring to Figure 14.1, it is seen that the voltage at point A is higher than the voltage at point B. If q was placed at d_1 and released, E would exert a force on it and move it toward d_2, that is, "downstream." From this, it can be concluded that *positive charges immersed in an* **E** field, if free to move, will migrate from a position of higher potential to a position of lower potential. Of course, a negative charge will migrate from lower to higher potential.

EXAMPLE 14.2

In Figure 14.1, $d_1 = 10.0$ cm, $d_2 = 40.0$ cm, and $E = 80.0$ N/C. Determine the voltage difference between points A and B.

Solution

Using Equations 14.8 and 14.9,

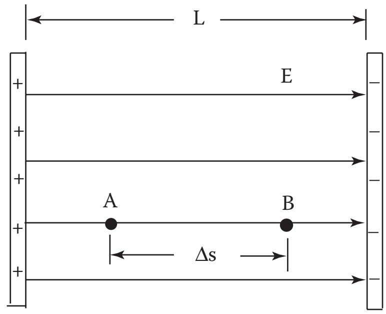
$$\Delta V_{AB} = \frac{\Delta PE}{q} = \frac{-W_{AB}}{q} = \frac{-Eq|d_2 - d_1|}{q} = -|E||d_2 - d_1|,$$

$$\Delta V_{AB} = V_2 - V_1 = -(80.0\ \text{N/C})[0.400\ \text{m} - 0.100\ \text{m}] = -24.0\ \text{V}.$$

Thus, point B is 24.0 V lower in potential than point A.

The change in potential ΔV may be related to the electric field directly. If an object of charge q is inserted into the uniform **E** field at point A (the following figure) and released, the conservative field will exert a force on it as it moves from point A to point B through a distance Δs. The work done by **E** is

$$W_{AB} = F_E \cdot \Delta s = -\Delta PE_{AB} = -q\Delta V \tag{14.10}$$



The work done by the electric field, on a charge, is proportional to the spatial change in electric potential.

or

$$(Eq)\Delta s = -q(\Delta V),$$

which gives

$$E = -\left(\frac{\Delta V}{\Delta s}\right). \tag{14.11}$$

The quantity $(\Delta V/\Delta s)$ is called the *potential gradient*. Although Equation 14.11 was derived for the case of a uniform field (the figure above), it also holds for a spatially nonuniform field. In that case, (Δs) is made infinitesimally small.

14.4 ELECTRIC POTENTIAL OF POINT CHARGES: SPATIALLY VARYING E

The spatially uniform E fields of Figures 14.1 and the figure under Example 14.2 led to interesting results that can, with proper mathematical techniques, be applied to nonuniform fields. The E field due to a positive point charge Q is

$$E_Q = \frac{kQ}{r^2} \tag{14.12}$$

and spatially decreases with increasing distance from Q. If a positive charge q is placed a distance r_P from Q and released, the field E_Q will repel it (Figure 14.2). The repulsive force is

$$F = \frac{kQq}{r^2}.$$

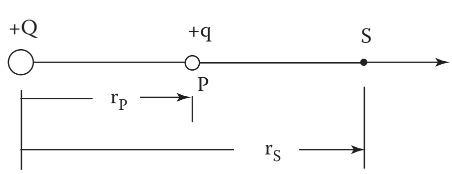


FIGURE 14.2 The electric field work done on a charge moved in a nonuniform field.

The work done by the field in moving q to the point r_S cannot be calculated from a simple ($\mathbf{F}\,\Delta\mathbf{r}$) equation since F varies with r, and hence is nonuniform. Integral calculus can be used to calculate W_{PS} and yields

$$W_{PS} = \frac{kQq}{r_P} - \frac{kQq}{r_S}.$$

Using Equation 14.10 gives

$$\Delta V = V_S - V_P = -\frac{W_{PS}}{q} = \frac{kQ}{r_S} - \frac{kQ}{r_P}. \tag{14.13}$$

Since $r_S > r_P$, ΔV is negative, that is, there is a decrease in potential from point P to point S. This corroborates the earlier statement that a (+q), if capable of movement, will migrate from a region of higher potential to a region of lower potential.

Electric potential is related to electric potential energy. Energy *differences* are the relevant quantities and a reference energy is considered. As an example, in the gravitational case, the surface of the earth is often assumed as the reference of gravitational potential energy and it is usually defined to have the value of zero.

For point charges, it is conventional to assign the potential energy and hence the potential to be zero at $r = \infty$. Equation 14.13 then, with $r_S = \infty$, so that $kQ/r_S = 0$, gives

$$V = \frac{kQ}{r}. \tag{14.14}$$

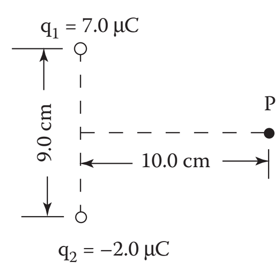
It is convenient to omit the subscript (P in this case) and recognize that Equation 14.14 gives the potential of the point charge Q at a distance r from Q. So Equation 14.14 gives the electric potential at location r relative to the potential at r equal to infinity.

Since potential is a scalar, the potential due to a group of charges can be found by simply adding the contribution due to each charge.

EXAMPLE 14.3

Two charges, $q_1 = +7.0\ \mu C$ and $q_2 = -2.0\ \mu C$, are 9.0 cm apart.

a. Determine the electric potential at point P which is 10.0 cm along the line that bisects the charge separation (the figure below).
b. The work required to bring a third charge $q_3 = +1.0\ \mu C$ from infinity up to point P.



The electric potential at point P due to two point charges q_1 and q_2.

Solution

a. $V = \sum_i \frac{kq_i}{r_i} = \frac{kq_1}{r_1} + \frac{kq_2}{r_2}.$

Here,

$$r_1 = r_2 = [(0.045\ \text{m})^2 + (0.10\ \text{m})^2]^{\frac{1}{2}} = 0.11\ \text{m}$$

so,

$$V = \frac{(9.0 \times 10^9\ \text{Nm}^2/\text{C}^2)}{(0.11\text{m})}[(7.0 \times 10^{-6}\text{C}) + (-2.0 \times 10^{-6}\text{C})]$$

or

$$V = 4.1 \times 10^5\ \text{V}.$$

b. $W = -\Delta PE = -q_3\Delta V = -(1.0 \times 10^{-6}\ \text{C})(4.1 \times 10^5\ \text{V})$

$$W = -0.41\ \text{J}.$$

The answer in part (b) of Example 14.3 was expressed in joules. When dealing with single charged particles, it is sometimes convenient to express their energy in terms of a unit called the *electron volt* (eV). The electron volt is defined as the amount of energy gained or lost by an electron when it traverses a potential difference of 1.00 V. Numerically,

$$1.00\ \text{eV} = q\Delta V = (1.00)(1.60 \times 10^{-19}\ \text{C})(1.00\ \text{J/C})$$

or

$$1.00\ \text{eV} = 1.60 \times 10^{-19}\ \text{J}.$$

14.4.1 Equipotential Surfaces

Surfaces, real or mentally constructed, on which the electric potential has the same value at all points are called *equipotential surfaces*. Some examples are as follows: (a) for a point charge, the equipotential surfaces are mentally constructed spherical surfaces of radius r = constant, centered on the charge; (b) for a uniform **E** field (Figure 14.1), they are planes parallel to the charged plates and perpendicular to **E**; (c) for a charged metal conductor with excess charge, its surface is equipotential; and (d) for a group of point charges,

$$\sum_i \frac{kq_i}{r_i} = \text{constant}$$

determines the equipotential surfaces. Two interesting and useful properties of equipotential surfaces are as follows:

1. No work is required to move a charge on the surface. Given

$$\Delta V = -\frac{W_{if}}{q} = V_f - V_i = 0,$$

so $W_{if} = 0$ since $V_f = V_i$.

2. The electric field on the surface is perpendicular to it and points in the direction of decreasing potential. This can be understood by noting that, since $W_{if} = 0$, there is no component of the electric field parallel to the surface, only a perpendicular component. Also, positive charges, free to move in an electric field, will migrate from higher potential to lower potential, that is, in the direction of decreasing potential.

PROBLEMS

14.1 Determine the absolute potential at a distance of 2.00 m from a + 3.00 μC charge.
Ans: 13,500 V.

14.2 The electric potential at a distance of 0.15 m from a point charge is found to +120 V. Determine the magnitude and sign of the charge.
Ans: 0.0020 μC.

14.3 An electrostatic force does 4.00×10^{-6} J of work in bringing a charge of $q = 3.00 \times 10^{-9}$ C from r = infinity up to some point P in space. Determine the electrostatic potential at P. (*Note:* V(infinity) = 0.)
Ans: 1.33×10^3 V.

14.4 A (–2.50 μC) charge of mass $m = 1.50 \times 10^{-6}$ kg is released from rest at point A, accelerates toward point B, and arrives at point B with a speed of 37.0 m/s. Determine
a. The numerical value of $(V_B - V_A)$.
b. The point at the higher potential.
Ans: (a) 411 V, (b) V_B.

14.5 Two charges, each of 2.50×10^{-5} C, are initially affixed to diagonally opposite corners of a square that is 0.300 m on a side. If one of the charges is moved to an empty corner of the square, how much work is done by the electric field?
Ans: –5.49 J.

14.6 Two protons are moving along a common straight line and headed directly toward each other. When they are very far apart, that is, infinitely separated, each has an initial velocity of 1.10×10^6 m/s. Determine their distance of closest approach.
Ans: 1.14×10^{-13} m.

14.7 An isolated (2.00×10^{-7} C) point charge is surrounded by mentally constructed spherical equipotential surfaces. Determine the value of the equipotential on a surface of radius r = 0.110 m from the charge.
Ans: 1.64×10^4 V.

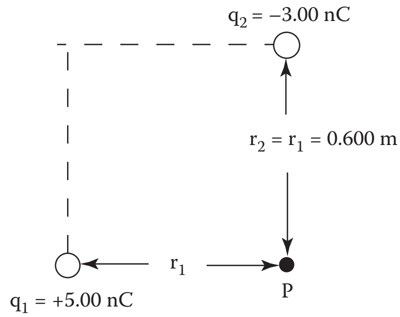
14.8 The gap (space between two metal conductors) on an automobile spark plug is set at 0.0450 in = 1.14 mm. When an electric spark jumps the gap to "fire" the fuel in the cylinder, the electric field in the gap is found to be 1.93×10^7 V/m. Determine the potential difference across the gap.
Ans: $\Delta V = 22{,}000$ V.

14.9 The outer surface of a cell membrane carries a positive charge and the inner surface carries a negative charge. The potential difference across the membrane is approximately 0.700 V and its thickness is 8.00×10^{-9} m. Determine the magnitude and direction of the electric field in the membrane.
Ans: $E = -8.75 \times 10$ V/m (huge); direction is from outer to inner surface.

14.10 An electron is moved from a point where the potential is (–75.0) V to a point where the potential is (+125.0) V. Determine the work done on the electron by the electric field.

14.11 Point charges of $q_1 = -11.0$ nC and $q_2 = -18.0$ nC are 0.900 m apart and fixed in space. Determine the electrostatic potential on a straight line joining the charges and at a point midway between them. *Note:* $1.0\text{ nC} = 1.0 \times 10^{-9}$ C.

14.12 A square, 0.600 m on a side, has two charges, $q_1 = +5.00$ nC and $q_2 = -3.00$ nC situated at opposite corners. Determine the electric potential at another corner P. There is no charge located at point P.



14.13 A charged particle (q = 2e), initially at rest, travels 1.20 cm due to a uniform electric field of magnitude 8.00×10^3 N/C. Determine the final kinetic energy of the charge (an alpha particle).

14.14 A proton is accelerated from rest through a potential difference ΔV. The proton attains a speed of 1.70×10^5 m/s. Determine the magnitude and "sign" of ΔV.

14.15 The molecules in air become ionized (dielectric breakdown) when in electric field approximately $\geq 3.00 \times 10^6$ V/m. Determine the maximum charge that can be placed on the surface of a metal conducting sphere of radius r = 10.0 cm without breakdown.

15 Direct Current Circuits

15.1 INTRODUCTION

Electric circuits are an excellent example of employing the properties of electric charges, covered in Chapters 13 and 14, for practical purposes. Circuits generally consist of (1) a source of charge, for example, a battery or an electric generator, (2) a conducting path for charge and hence energy flow, (3) a "load," that is, a charge/energy consuming device such as a radio, light bulb, and electric motor.

15.1.1 Cells and Batteries as a Source of Charge

A simple chemical cell can be constructed by placing two dissimilar metals, say A and B, in a dilute acid. As an example, the metals (electrodes), generally in the shape of rods, could be copper and zinc. Most metals dissociate (dissolve) slightly in the acid. The dissociation results in a positive ion being removed from the metal and entering the acid solution. That metal ion leaves behind *at least one* electron and therefore negatively charges the rod from which it came. Over time, as the process continues, the rod becomes negative enough to attract some of the positive ions in solution back to it. Eventually, an equilibrium is reached so that the number of positive ions leaving the rod equals the number attracted back to it. The electrode now is negatively charged, so a potential difference exists between it and the solution, which is at a higher potential. A similar process occurs with rod B and it likewise reaches an ion dissociation–attraction equilibrium. But metal B is dissimilar to electrode A and suppose it is more electronegative, that is, retains more electrons per dissociation than A. Both A and B are negative, so they are at a lower potential than the solution, but because of its greater electron retention, B is at a lower potential than A. This potential difference is traditionally called the *electromotive force* (a misnomer) or *emf* of the cell. Thus, electrode A will be the positive and B will be the negative terminals of the cell. If now, a conducting wire is attached to rods A and B, electrons will flow from the lower potential B, through the wire to the higher potential at A (Figure 15.1). The connecting wire, often connected to an electrical circuit, is usually placed outside the solution. As electrons leave B, they destroy B's dissociation–attraction equilibrium and allow more positive ions to leave B and go into solution. Additionally, as more electrons are deposited on A, they attract some of these ions which "plate out" on A. Thus, electrons migrate through the wire from B to A and an equal number of positive ions move through the solution from B to A. Therefore, the net charge on the electrodes remains unchanged. This process can continue until either B dissolves or its ions completely plate A. Then, the segments of the electrodes in solution are not "dissimilar." Charge ceases to flow. Note that the operation of commercially produced cells is somewhat more complicated but they operate by the same basic principles as discussed above.

The above discussion addresses a single cell with dissimilar metal electrodes A and B. The numerical value of the emf depends on the choice of metals and electrolyte (acid). Suppose a second similar cell with the positive electrode labeled A′ a conducting wire and negative electrode B′ is connected by a conducting wire to the first with B connected to A′ (Figure 15.2). The potential difference between A and B′ is now the *sum* of the potential differences of each cell separately. If many similar cells are connected in this fashion, which is called a *series* connection, the potential difference between the first positive and last negative terminals is the sum of the potential differences. Such a combination is called a *battery*. Your automobile battery consists of six cells each with a spongy lead (Pb) negative electrode and a lead oxide (PbO_2) positive electrode immersed in

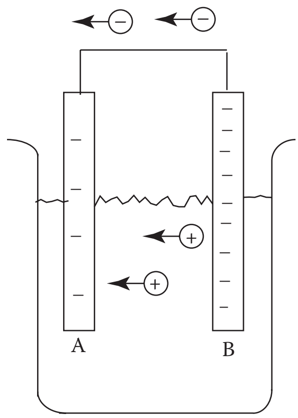


FIGURE 15.1 Two dissimilar metal rods placed in a dilute acid. A simple electric cell.

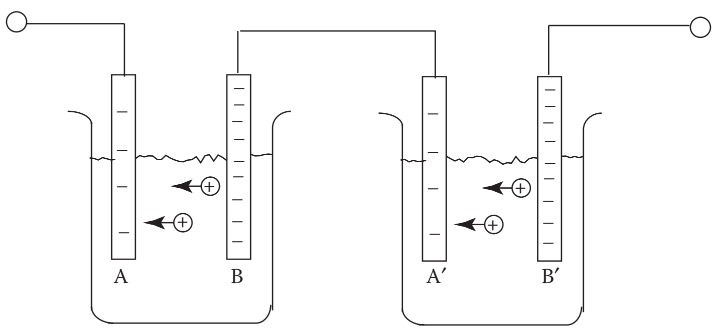


FIGURE 15.2 Two cells connected together to form an electric battery.

dilute sulfuric acid (H_2SO_4). The potential of each cell is 2 V and the series of six cells produces a battery of 12 V.

15.1.2 Electric Current

A potential difference and thus an electric field exist between the positive and negative terminals of a battery. If no conductor is connected to the terminals, no charge flows in the air between them. This is because the electric field between the terminals is generally not large enough to ionize the air molecules, which are charge neutral. However, if a conducting wire is connected to the electrodes, the battery produces an electric field within and parallel to the wire. This field exerts a force on the nearly free electrons in the wire and causes them to migrate from the lower potential negative terminal, to the higher potential positive terminal. This flow of charge constitutes an *electric current.*

The current is quantified by considering an imaginary surface, inside the conductor, whose plane is perpendicular to the direction of charge flow (Figure 15.3). The electric current is defined as the amount of charge that crosses this surface per unit time. If I represents the electric current, then

$$I = \frac{\Delta q}{\Delta t}. \tag{15.1}$$

The SI units of charge per time are coulomb per second (C/s). A current of 1 C/s is defined as 1 A (ampere) and is named in honor of the French mathematician Andre Ampere (1775–1836). If the current is not constant over the time interval (Δt), then Equation 15.1 gives the average current

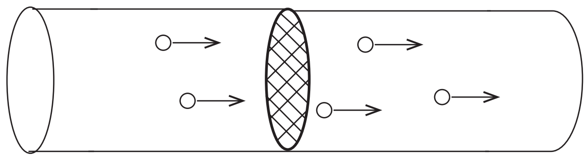


FIGURE 15.3 Moving charges in a conductor constitute an electric current.

during that interval. To give some perspective, we know that 1 C is equivalent to the charge on 6.25×10^{18} electrons, so 1 A implies 6.25×10^{18} electrons or protons crossing our imaginary surface per second.

The definition of electric current in Equation 15.1 applies to either positive or negative charges flowing in a conductor. In some electrolytic solutions, positive ions constitute the current, while, in some semiconductors, positive "holes" do. Early investigators of electricity believed that positive charges constituted the current in metal wires. For historical reasons, it is customary to use *conventional current*, that is, positive charge flow that would have the same electrical effect as electron flow. Conventional current flows from the positive terminal, through the circuit elements and into the negative terminal. This convention allows for a consistent use of algebraic signs.

15.2 OHM'S LAW

Ohm's law is a relation between electric potential difference and electric current. It is especially useful in analyzing the circuits. To make the discussion of circuits easier, it is customary to use symbols, not pictures to represent circuit elements. In Figure 15.4a, a cell or battery is depicted by a picture and the circuit symbol for it. Figure 15.4b shows a picture of a light bulb that takes energy from the charges flowing through it. Besides, it is the circuit symbol for the bulb or indeed any element that dissipates energy from the direct current source.

The circuit symbol for a conducting wire is a straight line. Figure 15.5 depicts a battery, connecting wires, and a "load," that is, a light bulb, direct current motor, or any energy-absorbing element. To establish Ohm's law, suppose that the cell in Figure 15.5 is a 1.5-V D cell of the kind used in an ordinary flashlight. Represent this voltage as $V_{1.5}$. For the given load R, the current is measured and found to be some value $I_{1.5}$. If now a second cell is connected in series to the first, the battery voltage applied to the circuit is doubled, that is, $V_{3.0} = 2V_{1.5}$. Again, the current is measured and found to be double its previous value, that is, $I_{3.0} = 2I_{1.5}$. If additional batteries of differing voltages are used to replace the original cell and the associated currents are measured, it is found that the current through the circuit, and hence the load R is directly proportional to the applied voltage. Thus I α V. This proportionality can be written as an equation if a multiplicative proportionality constant is inserted. Thus,

$$I = \left(\frac{1}{R}\right)V \tag{15.2a}$$

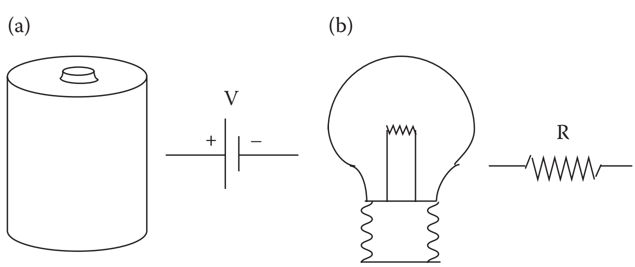


FIGURE 15.4 Some symbols used to represent circuit elements: (a) single cell and (b) resistor.

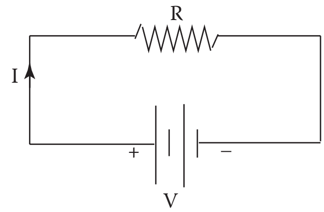


FIGURE 15.5 A simple electric circuit constructed of a source of charge, conducting wires, and a resistor.

or

$$V = IR. \tag{15.2b}$$

Equation 15.2b is called *Ohm's law* in honor of its discoverer, the German physicist Georg Ohm (1789–1854). The quantity R is called the *resistance*. This name is appropriate since Equation 15.2a indicates that, for a fixed voltage "V," the current I decreases as R increases. So R represents a hindrance to charge flow, that is, a "resistance." The SI unit of resistance, R = V/I is V/A = ohms and is usually represented by the symbol Ω. Thus, a circuit element has a resistance of 1 Ω if it carries a current of 1 A when a potential difference (voltage) of 1 V is applied across it.

EXAMPLE 15.1: SIMPLE FLASHLIGHT

The circuit in Figure 15.5 could represent a simple flashlight. The applied voltage is V = 3.0 V and comes from two "D" cells connected in series. The resistance of the flashlight bulb, when lit, is approximately 8.0 Ω. Determine the current through the bulb.

Solution

Ohm's law gives

$$I = \frac{V}{R} = \frac{3.0\ \mathrm{V}}{8.0\ \Omega} = 0.38\ \mathrm{A}.$$

15.3 RESISTIVITY

The hindrance to charge flow, that is, the resistance of conductors, semiconductors, superconductors, and insulators, depends on both the geometry and microscopic properties of the resistor. Experiments show that the resistance of a conductor is directly proportional to its length, that is, $R \propto L$. For example, this implies that if a 1.0-m length aluminum wire has a given resistance, a 2.0-m piece of the same diameter will have twice that value. It has also been shown experimentally that the resistance is inversely proportional to the cross-sectional area of the resistor, that is, $R \propto 1/A$. Length and cross-sectional area are geometrical factors. Also, two different materials, say copper and tungsten wires of the same length and area, will have different resistances because of the differences in their microscopic characteristics. These characteristics called the *resistivity* are represented by the Greek letter ρ (Table 15.1). Thus, the resistance of a conductor of length L, cross-sectional area A, and resistivity ρ can be written as

$$R = \rho\left(\frac{L}{A}\right). \tag{15.3}$$

TABLE 15.1
Resistivity[a] of Metallic Conductors at 20°C

Material	Resistivity (ρ) (Ωm)
Aluminum	2.65×10^{-8}
Copper	1.72×10^{-8}
Gold	2.44×10^{-8}
Iron (99.99% pure)	9.71×10^{-8}
Lead	20.65×10^{-8}
Mercury	9.58×10^{-7}
Silver	1.59×10^{-8}
Steel	10.4×10^{-8}
Various insulators	10^{7}–10^{17}

Source: *CRC Handbook of Chemistry and Physics*, 61st Edition, Boca Raton, FL, 1980–1981, pp. E-86, F-173. With permission.

[a] The exact value for ρ depends on the purity of the material.

Mathematically, ρ may be thought of as a proportionality parameter. The SI units of ρ are, from Equation 15.3,

$$\rho = \left(\frac{RA}{L}\right) \rightarrow \frac{\Omega m^2}{m} = \Omega m.$$

EXAMPLE 15.2

Calculate the resistance of 1 m of number 12 gauge copper wire. Number 12 is the size typically used in house wiring and has a diameter of 2.05 mm or a cross-sectional area of 3.31×10^{-6} m².

Solution

Using Equation 15.3

$$R = \rho\left(\frac{L}{A}\right) = (1.72 \times 10^{-8}\ \Omega m)\left(\frac{1.00\ m}{3.31 \times 10^{-6}\ m^2}\right) = 5.20 \times 10^{-3}\ \Omega.$$

A small resistance indeed. It would require 192 m of such wire to have a resistance of 1.00 Ω.

15.4 SIMPLE CIRCUITS

15.4.1 Series Connections

Often, circuits consist of more than one circuit element or device as indicated in Figure 15.6. As an example, R_1, R_2, and R_3 could be three bulbs in an in-store window display. The three devices are said to be connected in *series. A series connection is one where the same physical current passes through each device*. In Figure 15.6, when switch S is closed, the battery, with terminal potential V establishes a current I_T in the circuit. I_T leaves the (+) terminal, goes through resistors R_1, R_2, and R_3 and enters the (–) terminal. If the currents through the resistors are labeled I_1, I_2, and I_3, respectively, the series connections imply that $I_1 = I_2 = I_3$. Note that Ohm's law applies to each resistor and to the circuit as a whole.

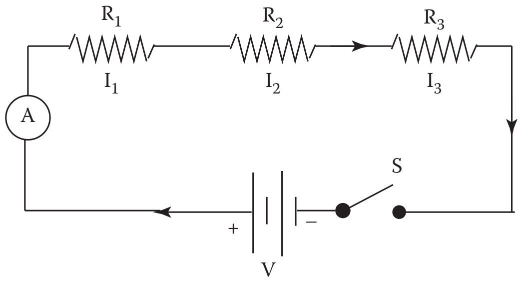


FIGURE 15.6 Three resistors connected in series with each other and a battery.

EXAMPLE 15.3

In Figure 15.6, $R_1 = 1.0\ \Omega$, $R_2 = 2.0\ \Omega$, $R_3 = 3.0\ \Omega$, and $V = 12.0$ V. Determine numerical values for (a) the total circuit resistance R_{TS}, (b) the total circuit current I_T, (c) the voltage drop across R_1, that is, V_1, (d) V_2, and (e) V_3.

Solution

a. The potential differences, that is, potential drops across each resistor must add up to the terminal voltage of the battery, that is,

$$V = V_1 + V_2 + V_3. \tag{15.4}$$

Inserting Ohm's law in Equation 15.4 gives

$$I_T R_{TS} = I_1 R_1 + I_2 R_2 + I_3 R_3. \tag{15.5}$$

Since $I_T = I_1 = I_2 = I_3$, Equation 15.5 becomes

$$R_{TS} = R_1 + R_2 + R_3 = 1.0\ \Omega + 2.0\ \Omega + 3.0\ \Omega = 6.0\ \Omega.$$

R_{TS} is called the *equivalent resistance* of the series circuit. This means that the three resistors R_1, R_2, and R_3 could be replaced by a single resistor R_{TS} that would have the same effect on the circuit as the three. It is always true that when resistors are connected in series, their equivalent resistance is the sum of the separate resistances. So,

$$R_{TS} = R_1 + R_2 + R_3 + \cdots. \tag{15.6}$$

b. Now that R_T has been determined, use Ohm's law for the circuit

$$I_T = \frac{V_T}{R_{TS}} = \frac{12.0\ \text{V}}{6.0\ \Omega} = 2.0\ \text{A}.$$

c. $V_1 = I_1 R_1 = I_{TS} R_1 = (2.0\ \text{A})(1.0\ \Omega) = 2.0$ V.
d. $V_2 = I_2 R_2 = (2.0\ \text{A})(2.0\ \Omega) = 4.0$ V.
e. $V_3 = I_3 R_3 = (2.0\ \text{A})(3.0\ \Omega) = 6.0$ V.

15.4.2 Parallel Connections

Another method of connecting circuit elements is in *parallel*. *A parallel connection is one where the same potential difference is applied across each device*. Figure 15.7 shows three resistors connected in parallel. The equivalent resistance for resistors connected in parallel can be found from recognizing that

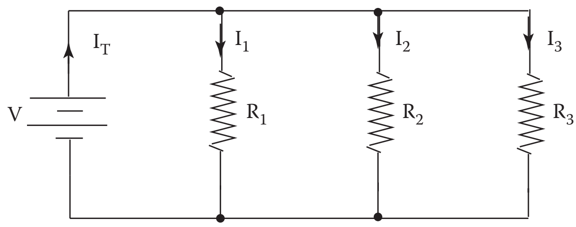


FIGURE 15.7 Resistors connected in parallel with each other.

$$I_T = I_1 + I_2 + I_3. \tag{15.7}$$

Ohm's law inserted into Equation 15.7 gives

$$\frac{V_T}{R_{TP}} = \frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}, \tag{15.8}$$

but $V = V_1 = V_2 = V_3$, so Equation 15.7 becomes

$$\frac{1}{R_{TP}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}. \tag{15.9}$$

Equation 15.9 can easily be extended to more than three devices.

EXAMPLE 15.4

In Figure 15.7, $R_1 = 1.0\ \Omega$, $R_2 = 2.0\ \Omega$, $R_3 = 3.0\ \Omega$, and $V_T = 12.0$ V. Determine numerical values for (a) total circuit resistance (equivalent resistance) R_{TP}, (b) circuit current I_T, (c) I_1, the current through R_1, (d) I_2, and (e) I_3.

Solution

a. Using Equation 15.8

$$\frac{1}{R_{TP}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} = \frac{1}{1.0\ \Omega} + \frac{1}{2.0\ \Omega} + \frac{1}{3.0\ \Omega} = \frac{6.0 + 3.0 + 2.0}{6.0\ \Omega} = \frac{11.0}{6.0\ \Omega}.$$

So,

$$R_{TP} = \frac{6.0\ \Omega}{11.0} = 0.55\ \Omega.$$

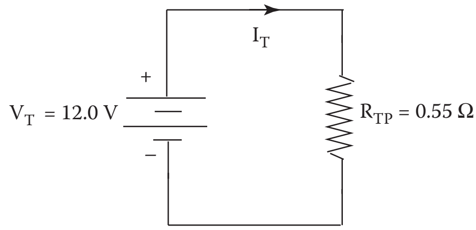
b. $I_T = \dfrac{V_T}{R_{TP}} = \dfrac{12.0\text{ V}}{0.55\ \Omega} = 22$ A.

c. $I_1 = \dfrac{V_1}{R_1} = \dfrac{V_T}{R_1} = \dfrac{12.0\text{ V}}{1.0\ \Omega} = 12$ A.

d. $I_2 = \dfrac{V_2}{R_2} = \dfrac{12.0\text{ V}}{2.0\ \Omega} = 6.0$ A.

e. $I_3 = \dfrac{V_3}{R_3} = \dfrac{12.0\text{ V}}{3.0\ \Omega} = 4.0$ A.

Thus, the circuit in Figure 15.7 can be replaced by the "equivalent" circuit of the following figure.



The equivalent circuit to Figure 15.7, calculated in Example 15.3.

15.4.3 Mixed Series–Parallel Connections

Circuits are often constructed consisting of some devices connected in parallel with each other. This parallel combination is often in series with other devices as in Figure 15.8. Analysis of such circuits is done by reducing the separate parts (series, parallel) to their equivalent values. This method is shown in the next example.

EXAMPLE 15.5

In Figure 15.8, $V_T = 8.0$ V. Determine values for (a) the equivalent resistance R_T, (b) the circuit current I_T, (c) the voltage drop V_1 across R_1, and (d) V_3.

Solution

a. The parallel combination can be replaced by a single equivalent resistor R_{TP} that is acquired by

$$\frac{1}{R_{TP}} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{20.0\ \Omega} + \frac{1}{30.0\ \Omega} = \frac{3+2}{60.0\ \Omega}$$

or

$$R_{TP} = \frac{60.0\ \Omega}{5} = 12.0\ \Omega.$$

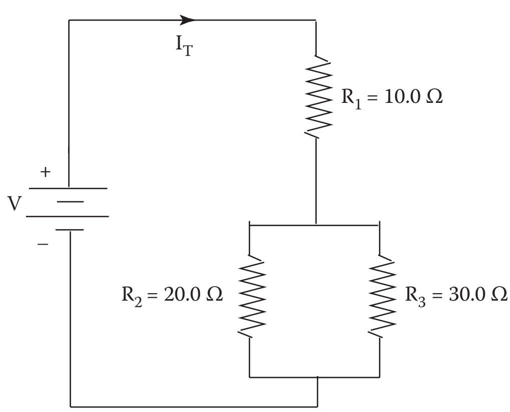
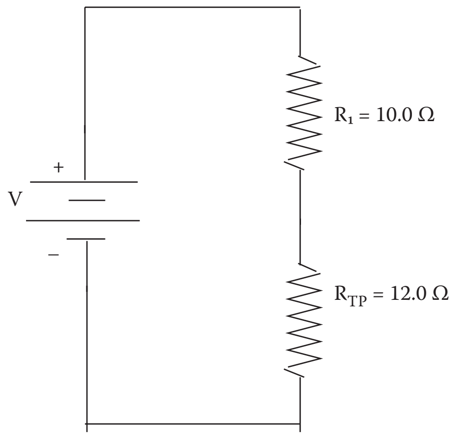


FIGURE 15.8 A combination series–parallel circuit.

The reduced circuit is shown in the figure below. Furthermore, R_1 and R_{TP} are in series and can be replaced by a single resistor R_T whose value is

$$R_T = R_1 + R_{TP} = 10.0\ \Omega + 12.0\ \Omega = 22.0\ \Omega.$$



The combination series–parallel circuit of Figure 15.8, reduced to a series circuit.

The reduced circuit is shown in the figure below.

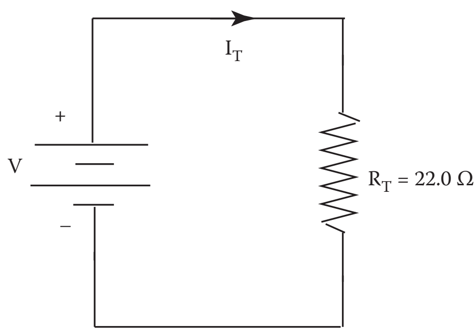


Figure 15.8 reduced to its simplest equivalent circuit.

b. The total current can be calculated by using the figure above. Its value is

$$I_T = \frac{V_T}{R_T} = \frac{8.0\ \text{V}}{22.0\ \Omega} = 0.36\ \text{A}.$$

c. From Figure 15.8, it is clear that the current through R_1 is I_T, so

$$V_1 = I_1R_1 = I_TR_1 = (0.36\ \text{A})(10.0\ \Omega) = 3.6\ \text{V}.$$

d. Since 3.6 V is dropped across R_1, the voltage drop across the parallel combination is

$$V_2 = V_3 = 8.0\ \text{V} - 3.6\ \text{V} = 4.4\ \text{V}.$$

15.5 ELECTRIC POWER

Sources of charge (emfs) do work on the charges they move through both attached resistors or complete circuits. The rate at which this work is done is the power provided by the source. So,

$$P = \frac{\Delta W}{\Delta t} = \frac{\Delta q V}{\Delta t} = \left(\frac{\Delta q}{\Delta t}\right)V = IV. \qquad (15.10)$$

Here, V is the terminal voltage of the source and the definition of electric current, $I = (\Delta q/\Delta t)$, has been used. The SI unit of power is the watt (W).

Equation 15.10 also applies to heat-dissipating devices such as toaster ovens and electric stove heating elements. The voltage in these cases is the voltage drop across such resistors. For resistors, Equation 15.10 can be written in two other forms. They are

$$P = IV = I(IV) = I^2R \qquad (15.11)$$

and

$$P = IV = \left(\frac{V}{R}\right)V = \frac{V^2}{R}. \qquad (15.12)$$

EXAMPLE 15.6

The two headlights of an automobile, if turned on, draw approximately 11.0 A when the voltage drop across them is 12.0 V. How much electrical power do they consume?

Solution

$$P = IV = (11.0\ \text{A})(12.0\ \text{V}) = 132\ \text{W}.$$

Note: 1.0 h.p. = 746 W, so the lights take approximately (132/746) = 0.177 h.p. to operate.

15.6 KIRCHHOFF'S RULES

Often, circuits are more complicated than the ones discussed above and are not amenable to either series or parallel reduction to equivalent circuits. For example, several charge sources (batteries) are placed between some of the resistors such as in Figure 15.9.

The analysis of such circuits is possible by use of *Kirchhoff's rules*, so named in honor of their originator, the German scientist Gustav Robert Kirchhoff (1824–1887). The rules are restatements of the laws of conservation of charge and conservation of energy applied to electric circuits. To apply Kirchhoff's rules, two associated definitions are needed:

i. *Junction*. A point in the circuit where three or more conductors are joined. In Figure 15.9, points b, c, f, and g are junctions.
ii. *Loop*. Any *closed* conducting path. In Figure 15.9, paths abgha, bcfgb, adeha, and so on are loops.

Kirchhoff's rules can now be stated.

I. *Junction rule*. The algebraic sum of currents at any junction must be equal to zero. Mathematically,

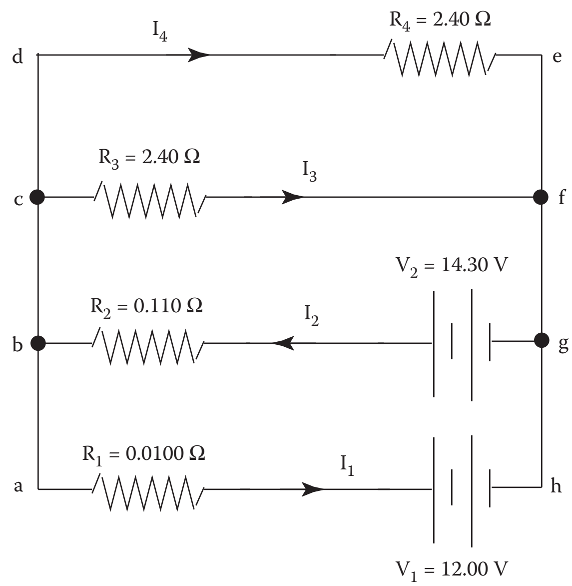


FIGURE 15.9 An example of a circuit not reducible to a simple equivalent circuit.

$$\sum_{\text{junction}} I = 0. \tag{15.13}$$

Sign convention: Currents are positive if entering a junction and negative if leaving, that is, directed away from the junction.

The junction rule is just a statement of the conservation of charge and implies that the junction is not a source or sink for charge. That is, charge does not "build up" or "disappear" at a junction.

II. *Loop rule*. Around any closed loop, the algebraic sum of the potential rises (emfs) equals the algebraic sum of the potential drops. Mathematically,

$$\sum_{\text{loop}} \varepsilon = \sum_{\text{loop}} (\text{P.D.}). \tag{15.14}$$

This rule is really a statement of the conservation of energy. It implies that, in traversing a closed circuital loop, any rises in potential and hence potential energy that a charge experiences must be balanced by an equal loss in potential. New symbols have been introduced in Equation 15.10. The εs are rises in potential and are associated with sources of charge, batteries, generators, and alternators, usually called emfs. The P.D.s are potential drops and are, up to now, associated with dissipative losses in resistors.

Sign convention

1. ε is a positive quantity in going from the negative to the positive terminal and is independent of the direction of the current.
2. Choose a direction for the current in each branch of the circuit. You may not know the correct direction so just pick one.
3. If the loop is traversed against the chosen direction of I, assign it a negative sign in the P.D. If traversed in the direction of I, give it a positive sign.

EXAMPLE 15.7

Figure 15.9 is a simplified version of the headlight circuit in a typical automobile. $V_1 = 12.0$ V is the battery. $V_2 = 14.30$ V is the alternator that both powers the headlights and charges the battery when the engine is running. $R_1 = 0.0100\ \Omega$ and $R_2 = 0.110\ \Omega$ are the internal resistances of the battery and alternator, respectively. $R_3 = R_4 = 2.40\ \Omega$ are the resistances of the headlight filaments when the lights are "on." Notice that they are connected in parallel, so if one "burns out" the other remains lit. Determine the currents I_1, I_2, and $I_3 = I_4$.

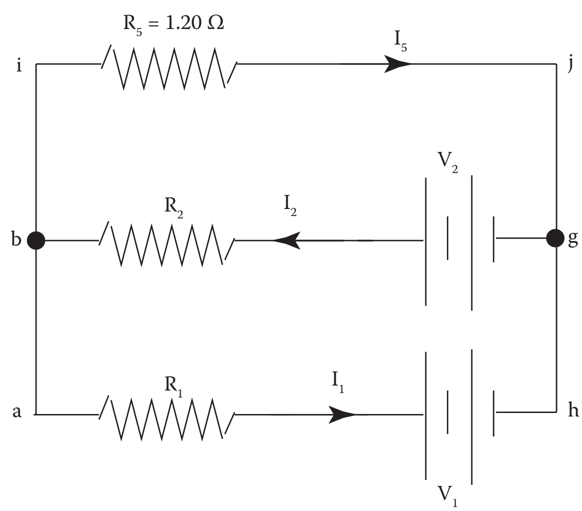
Solution

Since R_3 and R_4 are in parallel, with no battery in-between them, they can be replaced with their parallel combination R_5.

Here,

$$\frac{1}{R_5} = \frac{1}{R_3} + \frac{1}{R_4} = \frac{1}{2.40\ \Omega} + \frac{1}{2.40\ \Omega} = \frac{2}{2.40\ \Omega}.$$

So $R_5 = 1.20\ \Omega$. The circuit can be redrawn as in the figure below. Note that the directions of the currents are not generally known beforehand but are assumed. If a wrong direction is assigned, that current will have a negative sign in your answer. So, it does not matter if you initially choose the incorrect direction. There are three unknowns, I_1, I_2, and I_5, so three independent equations are required to solve uniquely for the currents.



A simplified automobile headlight circuit.

At junction b, rule (I) gives

$$\sum_b I = -I_1 + I_2 - I_5 = 0.$$

Traversing loop abgha clockwise, rule (II) gives

$$-V_2 + V_1 = (-I_1)R_1 + (-I_2)R_2.$$

Traversing loop bijgb clockwise, rule (II) gives

$$+V_2 = (+I_2)R_2 + (I_5)R_5.$$

Inserting values and rearranging the previous three equations gives

$$I_1 - I_2 + I_5 = 0,$$

$$(0.0100)I_1 + (0.110)I_2 = 14.30 - 12.00 = 2.30, \tag{15.15}$$

$$(0.110)I_2 + (1.20)I_5 = 14.30. \tag{15.16}$$

These three equations can be solved simultaneously for the currents. At this point, the manipulative steps for the solutions are a matter of personal taste. One possibility is to solve Equation 15.15 for I_1

$$I_1 = \frac{2.30 - (0.110)I_2}{0.0100} \text{ (}I_1 \text{ in terms of } I_2\text{)}.$$

Solve Equation 15.16 for I_5

$$I_5 = \frac{14.30 - (0.110)I_2}{1.20} \text{ (}I_5 \text{ in terms of } I_2\text{)}.$$

Inserting the above two equations into Equation 15.14 gives an equation with one unknown, I_2. Thus,

$$\frac{2.30 - (0.110)I_2}{0.0100} - I_2 + \frac{14.30 - (0.110)I_2}{1.20} = 0$$

or

$$230 - (11.0)I_2 - I_2 + 11.9 - (0.020)I_2 = 0.$$

Collecting like terms, $-(12.0)I_2 + 241.9 = 0$.
Solving,

$$I_2 = \frac{241.9}{12.0} = 20.2 \text{ A (alternator output current).}$$

This value of I_2 can be inserted into the equations for I_1 and I_5 in terms of I_2. They yield

$$I_1 = 8.25 \text{ A (charging the battery),}$$

$$I_5 = 11.9 \text{ A (powering the two headlights).}$$

Each headlight takes $(11.9)/2 = 6.0$ A.

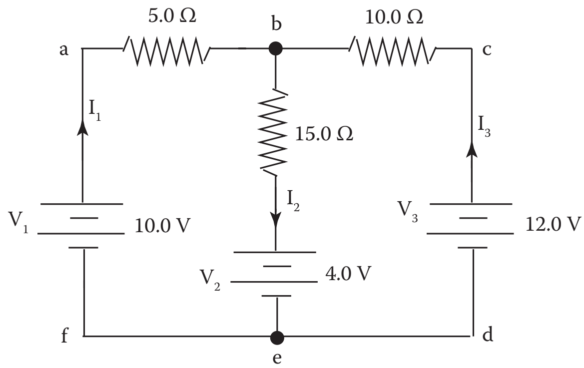
EXAMPLE 15.8

Determine all the currents and the voltage drop across the 5.0-Ω resistor for the circuit shown in the following figure. Which end of the 15.0-Ω resistor is at the higher potential?

Solution

Label the diagram and arbitrarily label and choose the current directions. Three currents imply that three independent equations are needed. At junction b,

$$I_1 - I_2 + I_3 = 0. \tag{15.17}$$



A circuit requiring Kirchhoff's rules for analysis.

Clockwise around loop abefa gives

$$+10.0\ \text{V} - 4.0\ \text{V} = (+I_1)(5.0\ \Omega) + (+I_2)(15.0\ \Omega). \tag{15.18}$$

Loop bcdeb counterclockwise gives

$$+12.0\ \text{V} - 4.0\ \text{V} = (+I_3)(10.0\ \Omega) + (+I_2)(15.0\ \Omega). \tag{15.19}$$

Again, the mechanics of manipulation to solve these three equations is a matter of choice. One possibility: solving Equation 15.18 for I_1 in terms of I_2 gives

$$I_1 = \frac{6.0\ \text{V} - (15.0)I_2}{5.0}. \tag{15.20}$$

Solving Equation 15.19 for I_3 gives

$$I_3 = \frac{8.0\ \text{V} - (15.0)I_2}{10.0}. \tag{15.21}$$

Insert Equations 15.20 and 15.21 into (15.17),

$$\frac{6.0\ \text{V} - (15.0)I_2}{5.0} - I_2 + \frac{8.0\ \text{V} - (15.0)I_2}{10.0} = 0. \tag{15.22}$$

Simplifying,

$$1.2 - (3.0)I_2 - I_2 + 0.80 - (1.5)I_2 = 0.$$

Collecting terms,

$$(5.5)I_2 = 2.0$$

or $I_2 = 0.36$ A. Now that I_2 has been determined, it can be used in Equations 15.20 and 15.21 to yield I_1 and I_3, respectively. Their values are

$$I_1 = 0.1\ \text{A} \quad \text{and} \quad I_2 = 0.3\ \text{A}.$$

The voltage drop across the 5.0-Ω resistor is

$$V = I_1R = (0.10)(5.0) = 0.5\ \text{V}.$$

Since I_2 is positive, its direction was chosen correctly. Current goes from higher to lower potential, the end closest to junction b is at the higher potential.

15.7 CAPACITORS

Capacitors are devices that store electric charge and hence electric field energy. They are typically constructed of two conductors separated by an insulating gap. The conductors may be of any shape or geometry. Flat plates and cylinders are the most common. Figure 15.10 depicts a simple series circuit with a parallel plate capacitor C.

Here, the plates are metal conductors and the insulating material in the gap is air. When switch S_1 is closed (conventional), charge leaves the positive terminal of the battery, migrates through switch S_1 and resistor R, and is deposited on plate A. An equal amount of positive charge leaves plate B and enters the negative terminal of the battery. Thus, plate B is equally negatively charged. This process continues, adding charge to plate A until it is at the same potential as the positive terminal of the battery. Charge ceases to flow. The capacitor is then "charged." If at any time during this charging process, the charge on either plate and the voltage across the plates are simultaneously measured, their ratio is found to be constant. This constant is called the *capacitance* C and is defined as

$$C \equiv \frac{|q|}{V}. \tag{15.23}$$

The SI unit of capacitance is the *farad* (F). From Equation 15.23,

$$1.00\ \text{F} \equiv \frac{1.00\ \text{C}}{1.00\ \text{V}}.$$

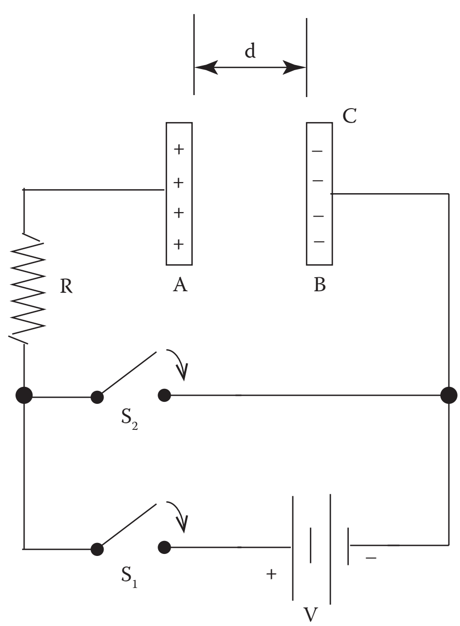


FIGURE 15.10 A simple parallel plate capacitor connected in a series circuit.

Equation 15.23 written as $q = CV$ implies that the charge stored on capacitor plates is directly proportional to the voltage drop across them and the proportionality parameter is the capacitance. It is not obvious from Equation 15.23, but capacitance is a geometric property of the plates and their separation d. It can be shown for a parallel plate capacitor where each plate is of area A and separation d, the capacitance is

$$C = \frac{\varepsilon_0 A}{d}. \tag{15.24}$$

If the space between the parallel plates is filled with a dielectric of dielectric constant κ, the capacitance is given by

$$C = \frac{\kappa \varepsilon_0 A}{d}. \tag{15.25}$$

The usual circuit symbol for a capacitor is ⊣ ⊢ but they are occasionally represented by the symbol —|(—.

The work done to charge a capacitor is equivalent to the energy stored between its plates. Recognize that as the source pushes additional charge onto the plates, it must do work against the charges it has already deposited there. The voltage across the plates thus varies from $V = 0$ (at the beginning of the charging process) to $V = V_f$ when fully charged. Since V is directly proportional to the charge on the plates, there is an average V, for the charging process, given by

$$V_{ave} = \frac{1}{2} V_f = \frac{1}{2} V. \tag{15.26}$$

So, the work required to charge the capacitor is

$$W = qV_{ave} = \frac{1}{2} qV \tag{15.27}$$

or energy stored is

$$E_{cap} = \frac{1}{2} qV = \frac{1}{2} CV^2 = \frac{1}{2}\frac{q^2}{C}, \tag{15.28}$$

where Equation 15.23 has been used to acquire the last two forms in Equation 15.28.

Referring to Figure 15.10, if S_1 remains closed, the capacitor will remain charged. If S_1 is opened, the charge remains on the plates. If now switch S_2 is closed (S_1 stays open), the positive charges on plate A will migrate around the circuit through S_2 and neutralize the negative charges on plate B. The capacitor has been "discharged."

15.7.1 Capacitors in Series

Figure 15.11 depicts three capacitors connected in series to a battery. Similar to the case for resistors, those three can be replaced by a single "equivalent" capacitor that will have the same circuital effect. When S is closed, Kirchhoff's loop rule gives

$$V = V_1 + V_2 + V_3.$$

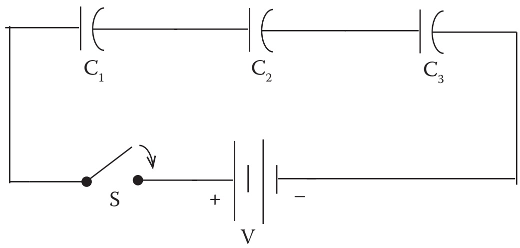


FIGURE 15.11 Capacitors connected in series with each other and a battery.

Inserting Equation 15.23 yields

$$\frac{q_T}{C_{TS}} = \frac{q_1}{C_1} + \frac{q_2}{C_2} + \frac{q_3}{C_3}.$$

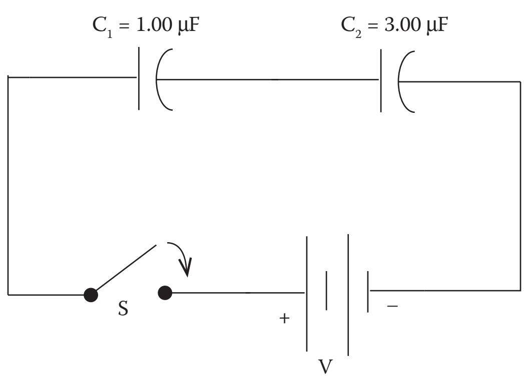
But $q_T = q_1 = q_2 = q_3$, so,

$$\frac{1}{C_{TS}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}. \tag{15.29}$$

The extension of Equation 15.29 to more than three capacitors, connected in series, is straightforward.

EXAMPLE 15.9

Determine the equivalent capacitance of the two series capacitors in the figure below.



The capacitive circuit that is analyzed in Example 15.9.

Solution

Using Equation 15.29

$$\frac{1}{C_{TS}} = \frac{1}{C_1} + \frac{1}{C_2} = \frac{1}{1.00\ \mu F} + \frac{1}{3.00\ \mu F} = \frac{3+1}{3.00\ \mu F}$$

or

$$C_{TS} = \frac{3.00\ \mu F}{4} = 0.75\ \mu F.$$

If the combination C_1 and C_2 were replaced by a single capacitor C_{TS}, the capacitive effect on the circuit would be the same.

15.7.2 Capacitors in Parallel

Figure 15.12 shows three capacitors connected in parallel to a battery. When switch S is closed, charge q leaves the battery. Some of that charge is deposited on C_3, some on C_2, and the remainder on C_1. So,

$$q = q_1 + q_2 + q_3.$$

Inserting Equation 15.23 gives

$$C_{TP}V = C_1V_1 + C_2V_2 + C_3V_3.$$

Since the capacitors are connected in parallel,

$$V = V_1 = V_2 = V_3.$$

Inserted this into the previous equation yields

$$C_{TP} = C_1 + C_2 + C_3. \tag{15.30}$$

Extension of Equation 15.30 to more capacitors connected in parallel is straightforward.

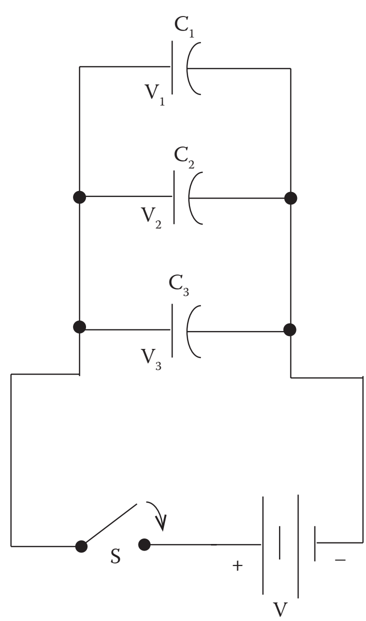


FIGURE 15.12 Capacitors connected in parallel with each other and with a battery.

EXAMPLE 15.10

Suppose the capacitors in Figure 15.12 have the values $C_1 = 1.00\ \mu F$, $C_2 = 2.00\ \mu F$, and $C_3 = 3.00\ \mu F$. Determine their equivalent capacitance.

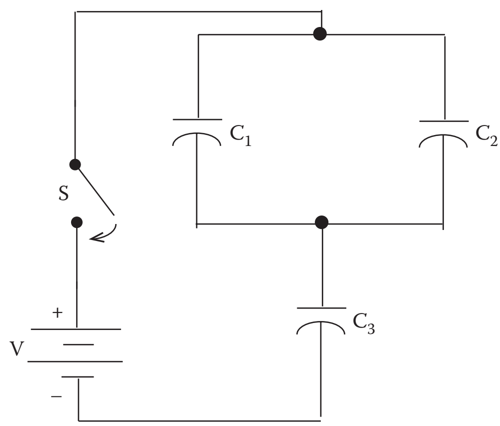
Solution

Using Equation 15.30

$$C_{TP} = C_1 + C_2 + C_3 = 1.00\ \mu F + 2.00\ \mu F + 3.00\ \mu F = 6.00\ \mu F.$$

EXAMPLE 15.11

Determine (a) the equivalent capacitance of the capacitors shown in the figure below, (b) the charge on C_3 when it is fully charged, (c) the energy stored in the equivalent capacitor, and (d) the energy stored in the fully charged C_2. Here $V = 12.0$ V, $C_1 = 4.00\ \mu F$, $C_2 = 5.00\ \mu F$, and $C_3 = 6.00\ \mu F$.



A more "complex" circuit with capacitors connected both in series and in parallel.

Solution

a. First reduce the parallel combination using

$$C_{TP} = C_1 + C_2 = 4.00\ \mu F + 5.00\ \mu F = 9.00\ \mu F.$$

C_{TP} is in series with C_3, so their equivalent capacitance is

$$\frac{1}{C_T} = \frac{1}{C_3} + \frac{1}{C_{TP}} = \frac{1}{6.00\ \mu F} + \frac{1}{9.00\ \mu F} = \frac{3+2}{18.00\ \mu F},$$

so,

$$C_T = \frac{18.00\ \mu F}{5} = 3.60\ \mu F.$$

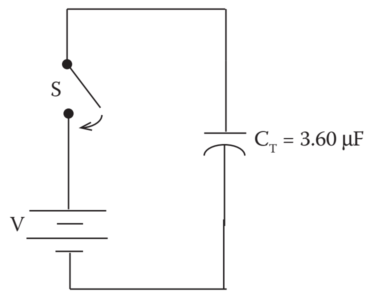
The equivalent circuit is shown in the following figure.

b. The total charge q_T on C_T is

$$q_T = C_T V = (3.60\ \mu F)(12.0\ V)$$

or

$$q_T = 43.2\ \mu C.$$



A simplified equivalent capacitive circuit.

The charge on C_3, q_3 is equal to q_T. So,

$$q_3 = 43.2\ \mu C.$$

c. The energy stored in C_T from Equation 15.28 is

$$E_{cap} = \frac{1}{2}\frac{q_T^2}{C_T} = \frac{1}{2}\frac{(4.32\times 10^{-5}\ C)^2}{(3.60\times 10^{-6}\ F)} = 2.63\times 10^{-4}\ J.$$

d. The value of C_2 is known, so by Equation 15.28 either q_2 or V_2 must be found to determine the energy stored in C_2. To find V_2, consider that

$$V_3 = \frac{q_3}{C_3} = \frac{q_T}{C_3} = \frac{43.2\ \mu C}{6.00\ \mu F} = 7.2\ V.$$

The parallel combination of C_1 and C_2 is in series with C_3, so

$$V_2 = V - V_3 = 12.0\ V - 07.2\ V = 4.8\ V,$$

then,

$$E_{cap} = \frac{1}{2}C_2(V_2)^2 = \frac{1}{2}(5.00\times 10^{-6}\ F)(4.8\ V)^2 = 5.8\times 10^{-5}\ J.$$

PROBLEMS

15.1 There is a current of 1.25 A through a light bulb. How many electrons flow through the bulb each second?

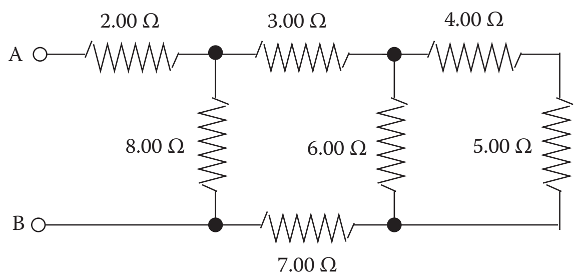
15.2 The filament of a "hot" light bulb has a resistance of 450 Ω and is connected across 120.0 V. Determine the current in the filament.

15.3 A new car battery is rated at 250 ampere hours (A h). This number indicates the total charge (under ideal conditions) the battery can "put out" before it is depleted. Determine
a. The total charge this battery can deliver to an automobile
b. The maximum current the battery can produce for 10.0 min
In reality, resistance and/or inductance would limit this current to about 400 A.

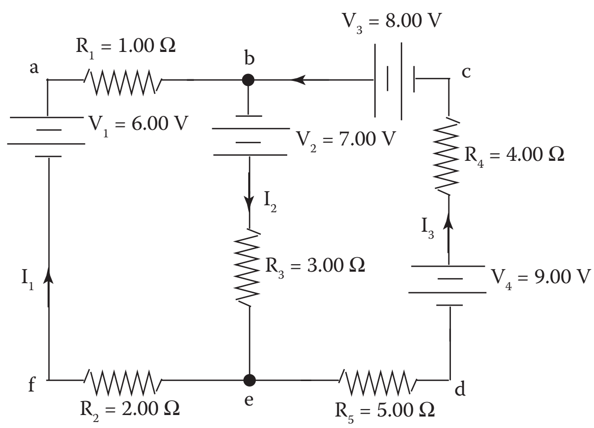
15.4 A long thin cylindrically shaped copper conductor carries a current of 15.0 A when a potential difference of 6.40×10^{-2} V exists across each meter of the wire. Determine the radius of this cable. *Note:* $\rho_{Cu} = 1.72 \times 10^{-8}\ \Omega m$.

15.5 Three resistors with values of 10.0, 7.0, and 3.0 Ω are connected in series across a 12.0-V battery. For each resistor, determine (a) its current, (b) voltage drop, and (c) electrical power dissipated.

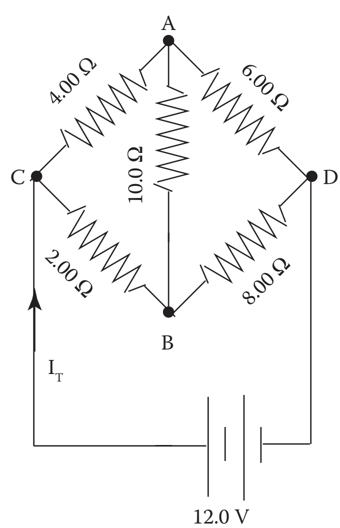
15.6 Determine the equivalent resistance between points A and B in the circuit shown below.



15.7 For the circuit shown below, $I_2 = 3.00$ A. Determine numerical values for I_1 and I_3.

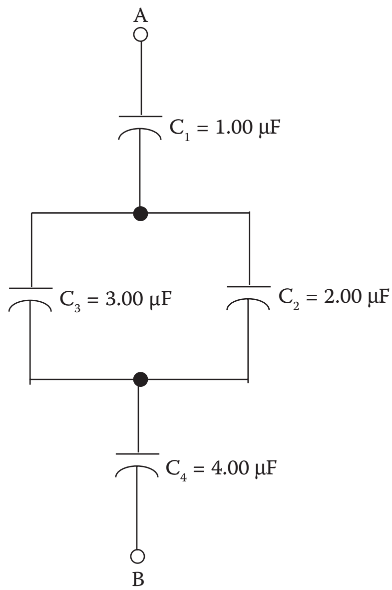


15.8 For the circuit shown below, determine the voltage between the points A and B. Which point, A or B, is at the higher potential? Establish your answer. *Note:* this is a challenging problem with five unknown currents and requires meticulous "bookkeeping."

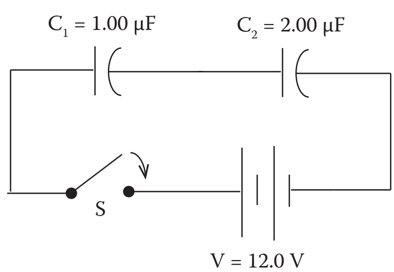


15.9 The heating coil of a water heater has a resistance of 15.0 Ω and operates at 220.0 V. Determine its power rating in watts.

15.10 Determine the equivalent capacitance between points A and B in the figure shown below.

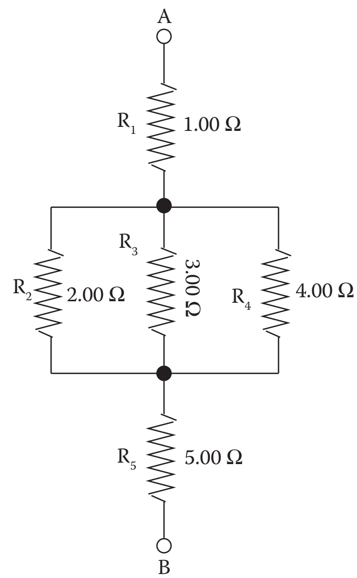


15.11 For the circuit shown below, determine (a) the equivalent capacitance, (b) the charge on capacitor C_2, (c) the voltage drop across C_2, and (d) the energy stored in C_2 when it is completely charged.



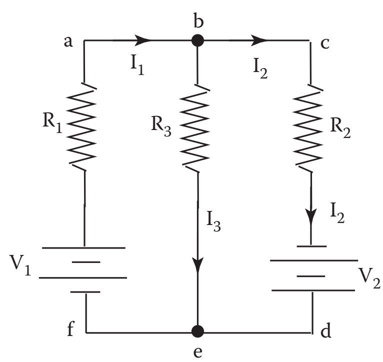
15.12 A cylindrical conducting wire has a diameter of 2.40 mm, a length of 4.00 m, and an electrical resistance of 0.433 Ω. Determine its resistivity ρ.

15.13 (a) Determine the equivalent resistance between points A and B in the following figure. (b) If a 12.0-V source is connected across points A and B, determine the voltage drop across resistor R_3.



15.14 An air-gap parallel plate capacitor has a plate area of 6.10×10^{-4} m^2 and a capacitance of 3.00 pF. Calculate the plate separation. *Note:* 1.0 picoFarad = 1.0 pF = 1.0×10^{-12} F.

15.15 Use Kirchhoff's rules to solve for the currents in each branch of the circuit shown below: $R_1 = 1.00\ \Omega$, $R_2 = 2.00\ \Omega$, $R_3 = 3.00\ \Omega$, $V_1 = 12.0$ V, and $V_2 = 9.00$ V.



16 Magnetic Forces and Fields

16.1 INTRODUCTION

Two empirical facts form the fundamental foundation of magnetism. They are as follows:

I. Charged particles, in motion in a magnetic field, experience forces of magnetic origin.
II. Moving charged particles create magnetic fields.

Historically, pieces of magnetite discovered in the province of Magnesia (Asia Minor) were found to exhibit an attraction or repulsion for each other. If two pieces of magnetite, that is, magnets were cut in the shape of rods or bars, it was found that they either attracted or repelled each other depending on the orientation of their ends. So, a force existed between them. Further, if one of these bars was placed in and attached to a small wooden boat situated in a tub of still water, the boat would slowly rotate and orient itself such that the long axis of the bar was approximately along a north–south line of longitude. The end of the bar oriented toward the earth's geographic north pole was called the north-seeking pole or, in short, the magnet's north pole. The opposite end of the bar magnet was called the south pole. Thus, the bar magnet could be used as a compass to point northward. If two of these bar magnets are brought close to each other such that the north pole of the first is closest to the north pole of the second, each experiences a force of repulsion (Figure 16.1a). If oriented such that the north pole of one and the south pole of the other are in closest proximity, each experiences a force of attraction (Figure 16.1b).

These magnetic forces exist if the magnets are situated in air, vacuum, dielectrics, and nonmagnetic conductors such as aluminum and copper. Forces transmitted through a vacuum, as in the case of electrostatic forces, represent action-at-a-distance forces. This implies a distortion of the space around the magnet, or a magnetic field.

16.2 MAGNETIC FIELD

Charged test particles are used to define and quantify the magnetic field. It is found that if a positively charged particle is placed between the poles of two anchored magnets, as in Figure 16.2a, no magnetic force is exerted on it. If the charged particle is given a velocity **v** as it passes through the region between the magnets, a magnetic force is exerted on it (Figure 16.2b).

This force is always found to be perpendicular to the particles velocity vector **v**. The magnetic field direction, symbol **B**, is defined as pointing from the north pole to the south pole of the separated magnet, as in Figure 16.3a. **B** and **v** thus form a plane. The magnetic force is found experimentally to be perpendicular to this plane and dependent on the magnitude of **v** and **B**. Given the directions of any two of the quantities **B**, **v,** and **F,** the direction of the third quantity can be found by use of the convenient mnemonic, the so-called right hand rule-I (RHR-I). To use RHR-I, position your right hand as shown in Figure 16.3b. Associate the direction of your extended fingers with the direction of B, the thumb should point along v and the direction in which the palm would push is the direction of **F**. Given any two of the directions, your right hand, with the above direction associations, determines the direction of the third.

Quantitatively, the magnetic field is defined as

$$B \equiv \frac{F}{qv \sin \theta}. \tag{16.1}$$

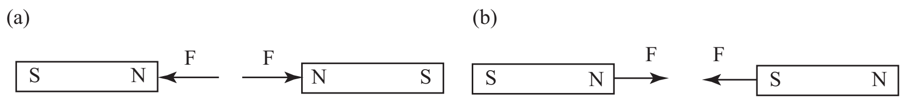


FIGURE 16.1 (a) Like magnetic poles repel each other. (b) Unlike magnetic poles exert a force of attraction on each other.

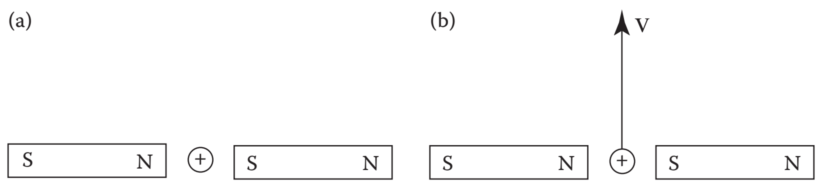


FIGURE 16.2 (a) A stationary charged particle, in a magnetic field, experiences no magnetic force. (b) A moving charged particle experiences a magnetic force.

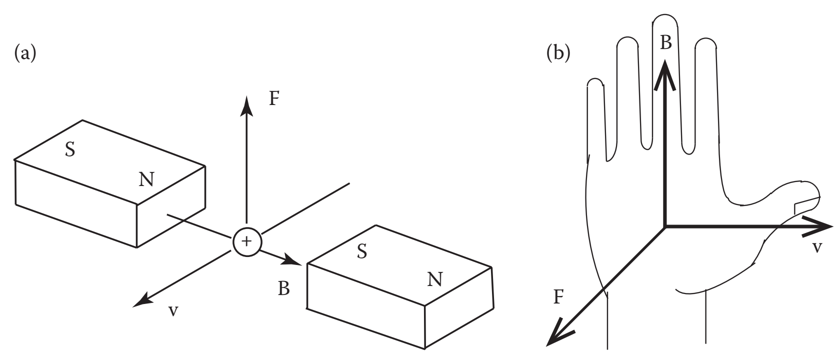


FIGURE 16.3 (a) The directions of the magnetic field, charged particle velocity, and magnetic force on the charge are mutually perpendicular. (b) RHR-I, a mnemonic for the relative directions of **B**, **F**, and **v**.

Here θ is the angle between the velocity vector **v** and the direction of the magnetic field **B**. Note that the maximum force on the moving charge occurs when **v** is perpendicular to **B**. The SI units of **B** are

$$B \rightarrow \frac{N}{C(m/s)} \rightarrow \text{Tesla (T)}.$$

So, if 1 C of charge with a velocity of 1 m/s experiences a magnetic force of 1 N, it is in a 1-T magnetic field. The unit is named in honor of the inventive Serbian-American scientist, Nikola Tesla.

If Equation 16.1 is rewritten as

$$F = Bqv \sin\theta. \tag{16.2}$$

It is clear that the magnetic force is directly proportional to the magnetic field, the magnitude and direction of the velocity, and the magnitude of the charge. Since F is perpendicular to v, the charge is deflected from a straight line path and follows a curved trajectory as shown in Figure 16.4.

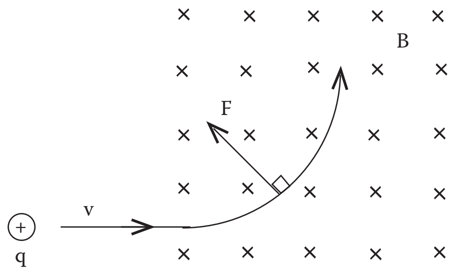


FIGURE 16.4 The curved trajectory of a charged particle moving through a magnetic field.

EXAMPLE 16.1

A charged particle of mass m, charge q, and velocity v is injected into a uniform magnetic field B as shown in the figure below. Determine an expression for the radius of its curved path in terms of m, v, q, and B.

Solution

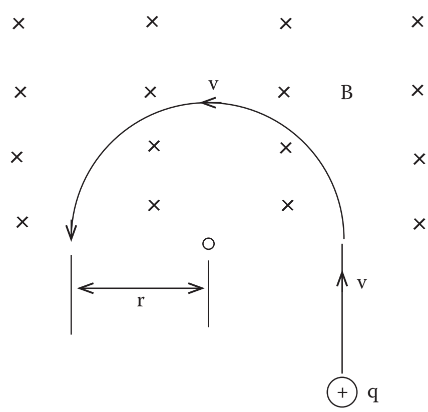
The force on q is centripetal, so Newton's second law gives

$$\sum F_B = ma_c = m\frac{v^2}{r}$$

or

$$Bqv \sin\theta = m\frac{v^2}{r}.$$

The velocity is in the plane of the paper, B is directed into the paper, so B and v are perpendicular throughout the semicircular path.



A simplified mass spectrograph.

So $\theta = 90°$, $\sin\theta = 1.0$, and

$$Bqv = m\frac{v^2}{r} \quad \text{or} \quad r = \frac{mv}{qB}. \tag{16.3}$$

Note that r is proportional to the particle mass. This is the basis of the mass spectrograph.

EXAMPLE 16.2: VELOCITY SELECTOR

It is possible to combine the use of both electric and magnetic fields to select or sort charged particles according to their velocity.

Solution

The figure below shows the cross section of two charged capacitor plates that create a downward pointing electric field. A positively charged particle moving between the plates would experience a downward force. The magnetic field is directed into the paper and would create an upward force on the charged particle. If these two forces are made equal in magnitude, the net vertical force will be zero and the particle will traverse the space undeflected. So,

$$\sum F_{vert} = 0 = F_B - F_E$$

or

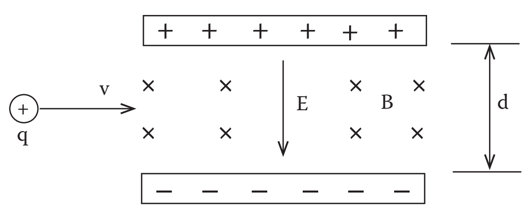
$$F_B = F_E$$

and

$$Bqv \sin\theta = qE, \quad \text{with} \ \ \theta = 90°,$$

so

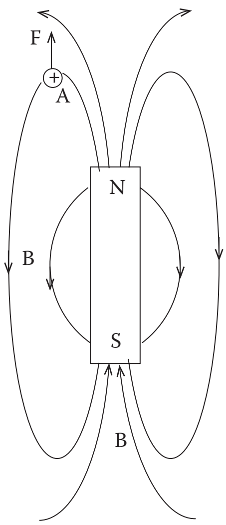
$$v = \left(\frac{E}{B}\right) = \frac{1}{B}\left(\frac{V}{d}\right). \tag{16.4}$$



Perpendicular E and B fields can be used to select a particular velocity from a beam of charged particles with various velocities.

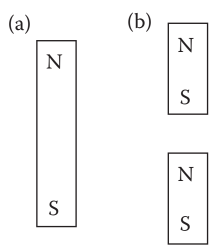
Here, V is the voltage across the capacitor plates and d is the plate separation. Note that Equation 16.4 selects a specific velocity. For particles with larger velocities, $F_B > F_E$; so they will hit the top plate.

For smaller velocities, $F_B < F_E$, and they will hit the bottom plate. A magnetic field exists in the space surrounding a bar magnet and its profile can be mapped by use of a moving positive test charge. The lines and arrows in the following figure represent the *direction* of the **B** field and *not* the direction of the magnetic force on a positive test charge. For example, a positive charge with velocity v directed into the paper at point A in the figure would experience an upward force as shown. Unlike the mapping of an electric field, where the lines of electric force are coincident with the lines of electric field, magnetic field lines are not coincident with the magnetic force on a moving charged particle. As stated earlier, the field lines are directed from the north pole of the magnet to its south pole.



The direction of the magnetic field around a bar magnet.

It is not possible to isolate a single north or south pole of a magnet. If the magnet (the figure below, part (a)) is cut into half, each half will possess a north and south pole (part (b)). The reason is explained in Section 16.5.



A single magnetic pole cannot be isolated.
They exist in N–S pairs.

16.3 MAGNETIC FORCE ON AN ELECTRIC CURRENT

Charged particles moving in a magnetic field experience magnetic forces. This is true whether the charges are free or constitute a current in a conductor.

Figure 16.5 depicts a current carrying segment of wire immersed in a uniform field. It has been shown experimentally that the wire experiences a magnetic force. To quantify this force, consider that at any moment an amount of conduction charge (Δq) is contained in a segment of wire of length (l) in the field. Equation 16.2 gives the magnitude of the force on these charges as

$$F = (\Delta q)vB \sin\theta. \tag{16.2}$$

Multiplying and dividing the right side by Δt gives

$$F = \left(\frac{\Delta q}{\Delta t}\right)(v\Delta t)B \sin\theta. \tag{16.5}$$

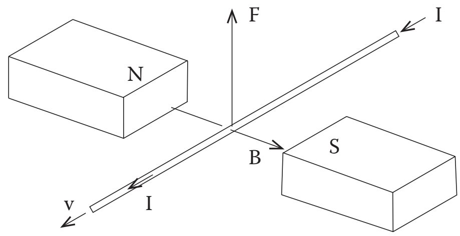


FIGURE 16.5 The force on a current-carrying conductor in a B field is perpendicular to both B and the direction of charge flow.

The quantity ($\Delta q/\Delta t$) is, by definition, the electric current and ($v\Delta t$) is the length of the segment of wire ℓ in the field. Equation 16.5 can be written as

$$F = BI\ell \sin\theta. \tag{16.6}$$

This important phenomenon, namely a current carrying conductor in a magnetic field experiences a magnetic force, has many practical applications. For example, it is the basis of operation for the electric motor and for most loudspeakers.

16.3.1 Torque on a Current Loop (Principle of the DC Electric Motor)

Consider a conducting wire with current I and shaped in the form of a rigid rectangle as shown in Figure 16.6. The rectangle, really an elementary armature, if placed in a magnetic field will experience magnetic forces on its four sides. The armature is fitted with supports and bearings (not shown) to allow it to rotate about an axis through its center and parallel to its length ℓ. Also affixed are two rigid conducting semicircular contacts called a commutator that rotates with the armature. The commutator makes contact with two stationary pieces of graphite, called brushes that permit electrical current to flow through the armature such that that segment of it that is in the right half-plane of Figure 16.7 always has its current flowing into the plane of the paper. Correspondingly, current flows out of the paper in the left half-plane segment. Figure 16.7 shows a cross section of the armature. The upper segment, with current directed into the paper, experiences a downward force. The bottom segment experiences an upward force. The two forces on these two segments produce a clockwise torque on the armature about its axis of rotation. The back segment experiences a force directed out of the paper and the force on the front segment is into the paper. These two forces "squeeze" on the armature but make no contribution to the torque about the axis of rotation. The magnitude of the torque due to the upper and lower segments is

$$\tau = [2(W/2)\sin\alpha]F_{mag} = W\left(BI\ell 1 \sin\theta\right)\sin\alpha \tag{16.7}$$

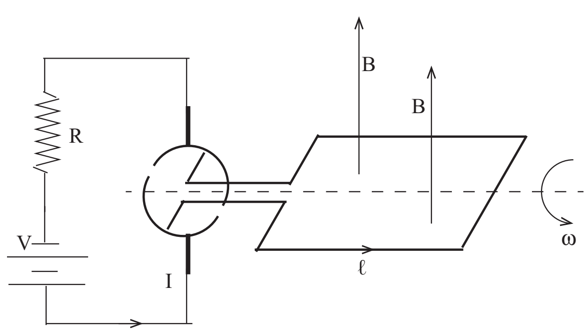


FIGURE 16.6 A current-carrying wire loop (armature), situated in a magnetic field, will experience a torque.

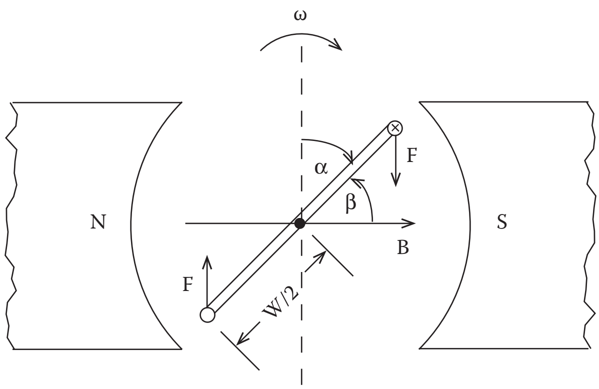


FIGURE 16.7 The geometry for analyzing the torque on an armature.

Here, θ is the angle between the directed line segments of the wires and the direction of the **B** field and is equal to 90° for segments, so sin θ = 1.0. Also, θ can be expressed in terms of the angular velocity θ of the armature as α = ωt. Equation 16.7 becomes

$$\tau = BI(W\ell)\sin \omega t. \tag{16.8}$$

The product ($W\ell$) = A is the area enclosed by the current loop. If N loops of wire are used instead of just one, each will experience a torque given by Equation 16.8. Thus, an armature consisting of N loops or "turns" experiences a torque given by

$$\tau = NBAI\sin \omega t. \tag{16.9}$$

So, send electric current through a correctly situated armature placed in a magnetic field and it will rotate continuously. We have a direct current (dc) motor.

EXAMPLE 16.3: SMALL DC MOTOR

A coil of wire (armature) has a length of 4.0 cm, width of 3.0 cm, and consists of 80.0 loops or "turns." It is placed in a uniform magnetic field of magnitude 0.12 T and given a current of 0.060 A. Determine the value of the maximum torque on the coil.

Solution

By Equation 16.9, maximum torque implies sin ωt = 1.0. The area of each loop is

$$A = \ell w = (4.0 \times 10^{-2}\ \text{m})(3.0 \times 10^{-2}\ \text{m}) = 1.2 \times 10^{-3}\ \text{m}^2.$$

Equation 16.9 gives

$$\tau = NBAI\sin \omega t$$

$$\tau = (80.0)(0.12\ \text{T})(1.2 \times 10^{-3}\ \text{m}^2)(6.0 \times 10^{-2}\ \text{A}) = 6.9 \times 10^{-4}\ \text{Nm}.$$

16.4 MAGNETIC FIELDS PRODUCED BY MOVING CHARGES

Charged particles, whether moving in air, a vacuum, through a conducting fluid, or as a current in a conductor, give rise to magnetic fields. Calculation of the magnetic field due to an isolated moving charge is beyond the scope of this chapter. However, the direction of the B field it produces can be determined by use of a convenient mnemonic called right hand rule-II (RHR-II) (Figure 16.8). To

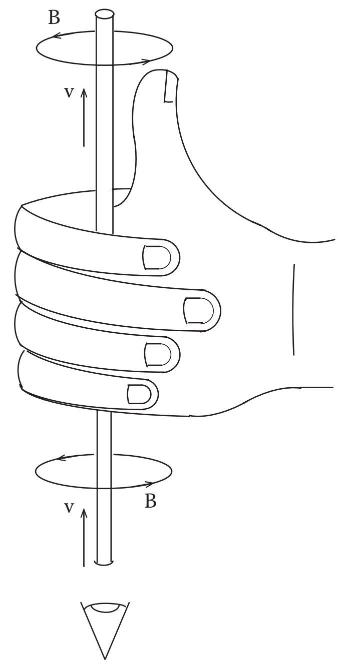


FIGURE 16.8 RHR-II gives the direction of a magnetic field due to moving charges.

execute this rule, for a positively charged particle do the following: (a) point the thumb of the right hand in the direction of the particle's velocity; (b) curl the fingers in a semicircle around the path of the moving charge, that is, clasp the particle path; and (c) the tips of the clasped fingers will point in the direction of the **B** field.

The **B** field will consist of concentric circles around the particle's path. The magnitude of **B** is found to decrease with increasing distance away from the particle path. For a positive particle, the concentric **B** field direction will be clockwise as one looks along the path in the direction of the velocity vector. Moving charges in a conductor, that is, electric currents, also create magnetic fields. The magnitude of these fields can often be calculated by either of two methods. They are (i) the Biot–Savart law and (ii) Ampere's law.

16.4.1 Biot–Savart Law

The Biot–Savart law is

$$|\Delta B| = \left(\frac{\mu_0 I}{4\pi}\right)\frac{|\Delta l||r_u|\sin\theta}{r^2}. \tag{16.10}$$

This equation implies that the current I in a small segment of wire (**Δl**) contributes an increment of magnetic field (**ΔB**) at a distance (r) to the space surrounding the wire.

The quantity r_u is a unit vector directed from the segment (**Δl**) to the spatial location of (**ΔB**) at r. The angle between (**Δl**) and r is θ. The quantity μ_0 is a proportionality constant called the *permeability of free space* and its value is $\mu_0 = 4\pi \times 10^{-7}$ Tm/A (Figure 16.9).

Equation 16.10 gives the contribution to the field due to a small current segment. To acquire the value of **B** at location r, the contributions due to all segments (**Δl**) must be summed over the total

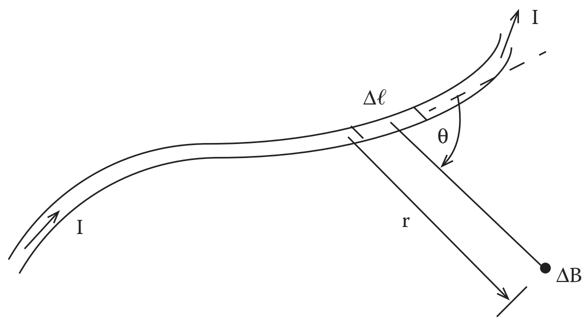


FIGURE 16.9 The current in a segment of a conductor contributes to the magnetic field at a distance r from the segment.

length of the conductor. Summing implies mathematical integration of equation (the Biot–Savart law 16.10). Depending on the shape or configuration of the conductor, the integration may not be simple. The results for some conductor geometries are listed as follows (Figure 16.10).

a. Long straight wire (the figure under Example 16.5):

$$|B| = \frac{\mu_0 I}{2\pi r},$$

where r is the distance from the wire.

b Circular loop at center (Figure 16.10):

$$|B| = \frac{\mu_0 I}{2r},$$

where r is the radius of the circular loop.

c. Solenoid of N turns (Figure 16.14):

$$|B| = \frac{\mu_0 NI}{L},$$

where L is the length of solenoid.

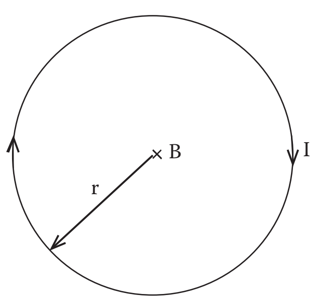
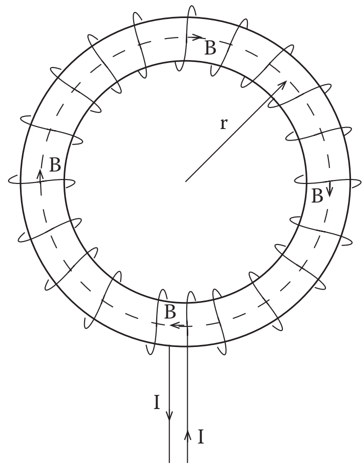


FIGURE 16.10 The illustration for Example 16.4.

d. Toroid



$$|B| = \frac{\mu_0 NI}{2\pi r}.$$

EXAMPLE 16.4

Determine the magnitude of the magnetic field at the center of a circular loop of wire of radius 5.00 cm and carrying a current of 20.0 A.

SOLUTION

Using the above equation for a circular loop,

$$|B| = \frac{\mu_0 I}{2r} = \frac{(4\pi \times 10^{-7}\ \text{Tm/A})(20.0\ \text{A})}{2(5.00 \times 10^{-2}\ \text{m})} = 2.51 \times 10^{-4}\ \text{T}.$$

16.4.2 Ampere's Law

Another, and perhaps more elegant, method of relating the electric current to the magnetic field that it creates is via Ampere's law. This law, discovered by the French physicist and mathematician Andre Ampere (1775–1836), is

$$\sum |B|(\Delta l)\cos\theta = \mu_0 I. \tag{16.11}$$

The use of Ampere's law involves forming a mental construct as was also the case for Gauss' law. Envision a mentally constructed closed loop that surrounds the conductor in a plane parallel to the conductor's cross section. Call this loop an *amperian loop*. The current in Equation 16.11 is the *net* current enclosed by or passing through this amperian loop. The **B** field in Equation 16.11 is evaluated at the location of the loop, $(\Delta \mathbf{l})$ is a segment of the loop, and θ is the angle between the direction of **B** and the directed loop segment $(\Delta \mathbf{l})$. The sum is performed over all segments of the closed loop. If $(\Delta \mathbf{l})$ is treated as an infinitesimal, the sum is replaced by an integral.

EXAMPLE 16.5

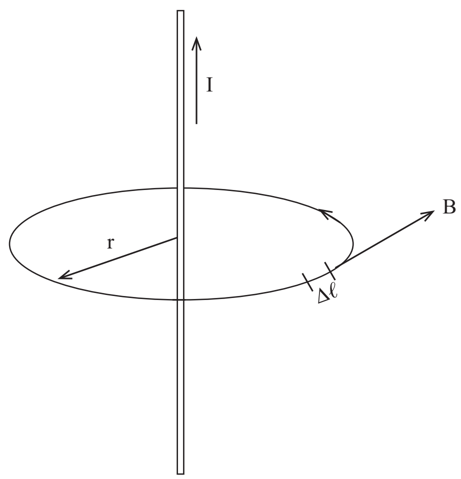
a. Use Ampere's law to derive an expression for the B field at a distance r from a long straight wire carrying a current I.
b. Estimate the fields numerical value if I = 25.0 A and r = 5.00 cm.

SOLUTION

a. Use Equation 16.11 and choose as the amperian loop a circle of radius r with the wire at its center. By RHR-II, the B field is directed as shown in the figure below. Perform the sum in Ampere's law around the amperian loop and choose (Δl) to be in the direction of **B** at all points on the loop. Then $\theta = 0°$ so $\cos\theta = 1.00$.

   By symmetry, the magnitude of **B** is the same at all points on the loop and can be factored out of the sum. So

$$B\sum(\Delta l) = \mu_0 I. \tag{16.12}$$



Geometry for the calculation of the magnetic field around a long, straight current-carrying wire.

The sum over (Δl) yields the circumference of the loop, that is, $\sum(\Delta l) = C = 2\pi r$ so Equation 16.12 gives

$$|B| = \frac{\mu_0}{2\pi}\frac{I}{r}.$$

b. Numerically,

$$|B| = \frac{(4\pi \times 10^{-7}\ \text{Tm/A})(25.0\ \text{A})}{2\pi(5.00 \times 10^{-2}\ \text{m})} = 1.00 \times 10^{-4}\ \text{T}.$$

This result is of the order of magnitude of three times the average value of the earth's magnetic field.

16.5 MAGNETIC MATERIALS AND PERMANENT MAGNETS

Current in a conducting circular loop gives rise to a **B** field. The field lines of such a loop, shown in Figure 16.11, are similar in shape to those of a bar magnet and indicate the directional sense of its north and south poles. If the loop had N turns, the field would have a larger value but the shape of the field lines would remain as shown.

It is conventional to indicate the field lines of Figure 16.11 by a single vector indicating the sense of the north pole (Figure 16.12). Additionally, a quantity called the *magnetic moment* (μ) is defined for the loop. Its *magnitude* is

$$\mu = NIA. \tag{16.13}$$

Here A is the area circumscribed by the loop and N the number of turns. The direction of μ is perpendicular to the plane of the loop and parallel to the vector direction of **A.** The loop is also called a *magnetic dipole*. If the loop is free to rotate and is placed in an *external* field (B_{ext}), a torque will be exerted on it. The loop will rotate so that μ is parallel to B_{ext} (Figure 16.13). This is a configuration of minimum energy. It can be shown that potential energy of such a loop in $\mathbf{B}_{ext}$ is

$$U = -\mu\left|B_{ext}\right|\cos\theta. \tag{16.14}$$

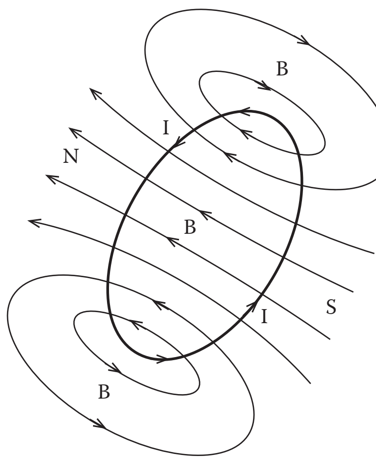


FIGURE 16.11 Magnetic field lines of a circular current loop.

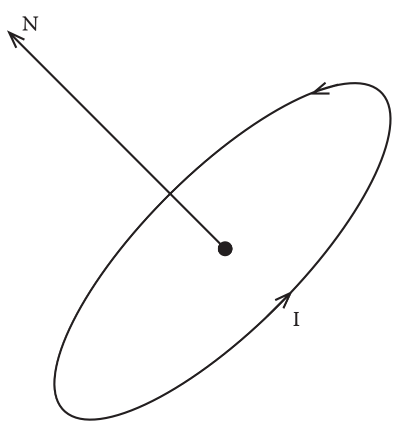


FIGURE 16.12 A simplified representation of the B field due to a current loop.

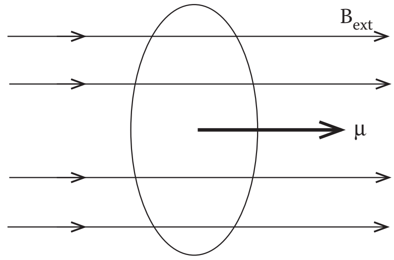


FIGURE 16.13 A current loop, in a magnetic field, will experience a magnetic torque.

The torque exerted on it is

$$\tau = |\mathrm{m}||\mathbf{B}_{ext}|\sin\theta. \tag{16.15}$$

Here, θ is the angle between μ and $\mathbf{B}_{ext}$. In an atom, via the simple planetary model, the motion of the electrons in orbit about the nucleus is similar to a current loop. Each electron thus gives rise to an orbital magnetic moment. If the atom has an odd number of electrons, total cancelation is not possible so it will have a net orbital magnetic moment (μ_l).

Additionally, it is found from the analysis of atomic spectra that electrons have intrinsic "spin" that, via a simple model, may be viewed as rotation about an axis through the electron. This spin has an associated spin magnetic moment (μ_s). Electrons tend to fill the atom shells in pairs, so often their spin moments (μ_s) tend to be oppositely directed and vectorally cancel each other. But, if the atom contains one or more unpaired electrons, it will have a net spin magnetic moment. Lastly, both the protons and neutrons possess intrinsic spin and therefore magnetic moments. These add vectorally to form a net nuclear magnetic moment (μ_N). The nuclear moment μ_N is approximately three orders of magnitude smaller than μ_l and μ_s. The net orbital μ_l, spin μ_s, and nuclear μ_N moments add vectorally to give a total magnetic moment (μ) for the atom. Thus, some atoms whose individual moments do not cancel have a net nonzero magnetic moment. These atoms behave as microscopic bar magnets. Magnetic properties of materials are characterized by the magnetic properties of their atoms. Materials whose atoms individually have no net magnetic moment, that is, $\mu = 0$ are called *diamagnetic*. Those whose atoms have a small net moment are *paramagnetic*. Materials whose atoms have a large μ and whose electrons possess a cooperative interaction with the electrons of neighboring atoms (the *exchange* interaction) are called *ferromagnetic*. The response of these three types of materials to an external magnetic field is characterized by their *relative permeability* $K_{m.}$

Consider a long solenoid of N turns wound in the form of a cylinder (Figure 16.14) and carrying a current I. The approximately uniform magnetic field $\mathbf{B}_0$ in the solenoid interior is given by

$$B_0 = \frac{\mu_0 NI}{L}.$$

Now insert magnetic materials into the interior of the solenoid and measure the fields interior to them (**B**). The ratio $K_m \equiv \mathbf{B}/\mathbf{B}_0$ gives the following for each material: diamagnetic $K_m < 1$, paramagnetic $K_m \geq 1$, and ferromagnetic $K_m >>> 1$.

Diamagnetism, $K_m < 1$: Diamagnetism is apparently caused by the action of the applied field $\mathbf{B}_0$ on the motion of the material's electrons. The result is a very small induced field that is oppositely directed to $\mathbf{B}_0$. Due to the fundamental nature of this interaction, all materials are thought to possess this very small effect. It can only be measured in materials that are neither paramagnetic nor ferromagnetic.

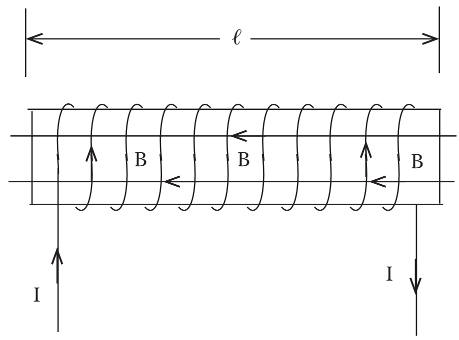


FIGURE 16.14 The magnetic field inside a long solenoid.

Paramagnetism, $K_m > 1$: The net magnetic moments of the electrons in these materials are nonzero but are randomly oriented. A torque, given by Equation 16.15, is exerted on these atoms when placed in an external field $\mathbf{B}_0$. This torque tends to partially align the net moments along the direction of $\mathbf{B}_0$ and hence vectorally add to it. Thus $\mathbf{B} > \mathbf{B}_0$.

Ferromagnetism, $K_m >>> 1$: The magnetic properties of ferromagnetic materials, such as iron, nickel, cobalt, gadolinium, and some of their alloys, are due to their unique electron configuration. Each has several electrons that are not paired with others of opposite spin. These atoms have a large spin magnetic moment. Because of the cooperative interaction between spins (exchange), large numbers of coupled atoms (10^{15}–10^{19}) have their moments pointing in the same direction. These moments add vectorally and create a small region that behaves as a permanent magnet. These small regions ≈$(0.05\ \text{mm})^2$ are called *magnetic domains*. The directions of the net domain moments are generally randomly oriented among the domains. In this case, the material does not behave as a magnet. If the material is placed in an external magnetic field, a torque is exerted on the domain moments. In some domains, the spins will "flip" in the direction of $\mathbf{B}_{ext}$ and add to it while in others, with a component of their spins along $\mathbf{B}_{ext}$, will expand at the expense of their neighboring domains. Which of these processes occur depends upon the minimization of internal crystal energies (domain wall movement, interatomic elastic energy, exchange energy, etc.). Both of these processes will greatly enhance the magnetic field in and around the material.

So a ferromagnet placed in a current carrying solenoid will create a much stronger magnetic field (Figure 16.15). This is the principle of the electromagnet.

If a ferromagnetic material is heated and melted, then allowed to cool and crystallize while simultaneously immersed in an external magnetic field, it will become a *permanent magnet* with a single domain (Figure 16.16).

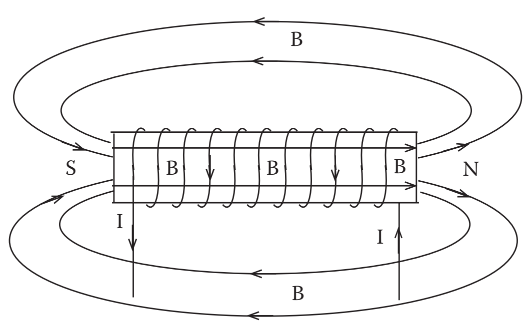


FIGURE 16.15 The magnetic field surrounding a ferromagnet enclosed in a solenoid.

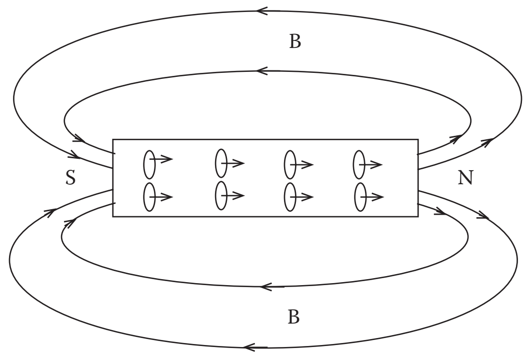
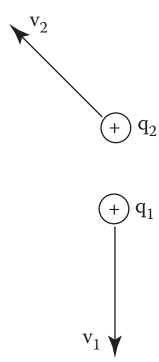


FIGURE 16.16 The magnetic field surrounding a magnetized ferromagnetic bar magnet.

PROBLEMS

16.1 A point charge $q = 1.60 \times 10^{-6}$ C experiences a magnetic force of 7.50×10^{-3} N when it moves through a magnetic field with a velocity of 2.50×10^4 m/s. The angle between the field direction and the charge's velocity vector is 40.0°. Determine the magnitude of the magnetic field.

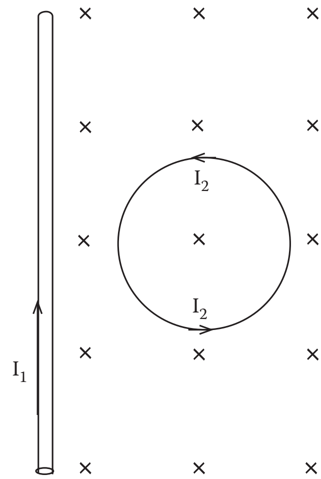
16.2 Two positive particles, q_1 and q_2, have velocities v_1 and v_2 in the plane of the page and directed as shown. The magnetic force on q_1 due to the motion of q_2, at the instant shown, is in what direction?
a. Vertically downward
b. Vertically upward
c. Horizontally to the left
d. Horizontally to the right
e. Out of the page and perpendicular to its plane
f. Into the page and perpendicular to its plane



16.3 An experiment is conducted whereby a particle of charge $q = 1.60 \times 10^{-19}$ C is successively injected into a 1.40-T field. The speed of the particle is 3.00×10^6 m/s. Determine the maximum force on the particle.

16.4 A long wire and a circular conducting loop carry currents I_1 and I_2, respectively. The wire and loop are in the plane of the paper with current directions as shown. The direction of the NET force on the loop, due to the currents is
a. Into the paper
b. Out of the paper
c. To the right, away from the wire

d. To the left, toward the wire
e. Vertically upward along the direction of I_1
Establish your answer.



16.5 A circular loop of radius r = 3.00 cm carries a current of 6.00 A. Determine the magnitude of the magnetic field at its center.

16.6 The magnetic field at the center of a long solenoid that is carrying a current of 8.00 A has the value $B = 1.21 \times 10^{-2}$ T. Determine the number of turns per meter for this solenoid.

16.7 The armature of a dc motor consists of 85.0 identical rectangular turns of wire, each of length 18.0 cm and width 10.0 cm. The armature is placed in a magnetic field that has an average value of 0.120 T. How much current must be supplied to this armature if it is to develop a torque of 6.00 Nm?

16.8 A very long straight cylindrical wire has a radius a = 3.00 mm and carries a current of 30.0 A. The current is assumed to be uniformly distributed over the cross section of the wire. Use Ampere's law to determine the magnetic field:
a. Inside the wire at r = 1.50 mm
b. Outside the wire at r = 7.00 mm

16.9 Use Ampere's law to derive an expression for the magnetic field at a distance "a" from a long straight wire carrying a current I. (*Hint*: Draw a circular amperian loop of radius "a," centered on the wire with plane perpendicular to it.)

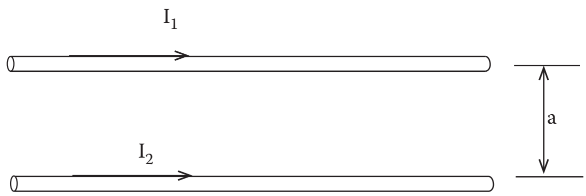
16.10 A point charge q = 1.80 μC is injected into a (0.300-T) uniform magnetic field with a velocity of 2.25×10^4 m/s and at an angle of 35.0° relative to the magnetic field direction. Determine the magnitude of the force on the charge.

16.11 An electron with velocity $v = 4.00 \times 10^6$ m/s is injected perpendicularly into a (0.250×10^{-3} T) uniform magnetic field. Calculate the radius of its curved semicircular path.

16.12 A velocity selector is to be adjusted to select charged particles with velocities of $v = 2.00 \times 10^5$ m/s. If the electric field is set at E = 10,000 V/m, to what magnitude should the magnetic field be adjusted?

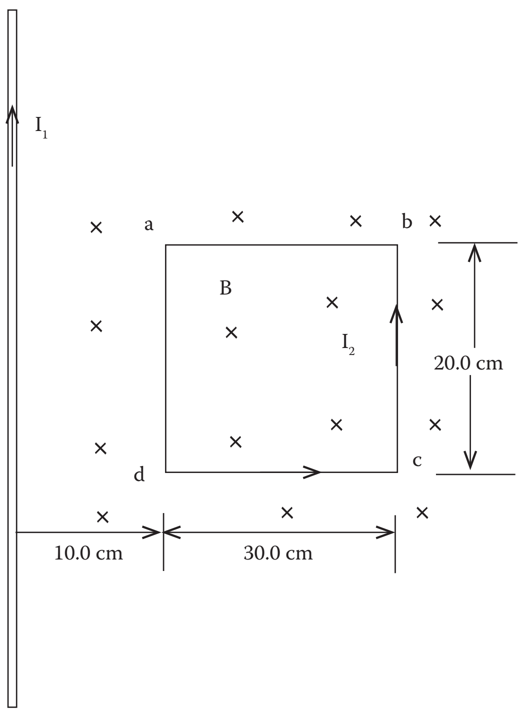
16.13 A long straight wire is oriented east–west in a region where the horizontal component of the earth's magnetic field is (0.177 T) and directed to north. The magnetic force on unit length of wire, perpendicular to the earth's surface, is found to be 0.00800 N/m. Determine the current in the wire.

16.14 Two long straight wires are both in the plane of the page and carry equal currents $I_1 = I_2 = 3.00$ A. They are separated by a distance a = 4.00 cm as shown. Determine (a) the magnitude of the magnetic field at wire "two" due to the current I_1; (b) the force per length of wire, exerted on each of the wires; and (c) is the force attractive or repulsive?



16.15 The armature of a dc motor consists of 100 identical square turns of wire. It is placed in a magnetic field of average value (0.130 T). The motor should produce a torque of 7.00 Nm when drawing a current of 15.0 A. (a) What value of area should be circumscribed by the turns? (b) What will be the length of the armature?

16.16 A rectangular loop, carrying a current $I_2 = 15.0$ A, is located 10.0 cm from a long, straight wire with current $I_1 = 35.0$ A. The dimensions of the loop are 30.0 cm by 20.0 cm. Calculate the approximate force on the rectangle due to the current in the long wire. (*Hint*: As a first approximation, treat each of the four segments of the rectangle, as long straight wires.)



16.17 Use the Biot–Savart law to acquire an expression for the magnetic field at the center of a circular conductor of radius r and carrying a current I.

17 Electromagnetic Induction

17.1 INTRODUCTION

Electromagnetic induction, that is, the generation of induced electromotive forces (emfs) and their associated currents (Faraday's law) is a beautiful law of physics and is an excellent illustration of the aesthetic behavior of nature. The practical consequences of it are the basis for the electrification of modern civilization. The brilliant English experimentalist, Michael Faraday (1791–1867), and an equally competent American scientist, Joseph Henry (1797–1878), performed the first recorded experiments on induction (Figure 17.1). Faraday published his experiments earlier than Henry so his name is associated with the phenomenon and the law that quantifies it. The foundations of Faraday's law are illustrated by the following experiments.

Experiment 1: A bar magnet thrust into the region circumscribed by a conducting closed loop induces an emf and associated current when the magnet is in motion (Figure 17.1). When the motion ceases, the emf around the circuit is zero. If the magnet is now withdrawn, a current is induced in the opposite direction during the motion. The polarity of the emf and the direction of the current are indicated by the galvanometer. So, the motion of the bar magnet, when it is moved back and forth through the loop, induces a voltage in the loop. Of course, if the magnet is held stationary and instead the rigid wire loop is thrust back and forth around the magnet, an emf will be induced in the loop also. Additionally, if both the magnet and loop are stationary but the loop is distorted, that is, stretched or shrunk such that it circumscribes a changing area, an emf will be induced in the loop only during the time when the area is changing. Finally, if the loop is rotated about an axis, while the magnet is stationary, an emf will be induced in it. In this case, the loop would need to be connected to the galvanometer by sliding circular contacts.

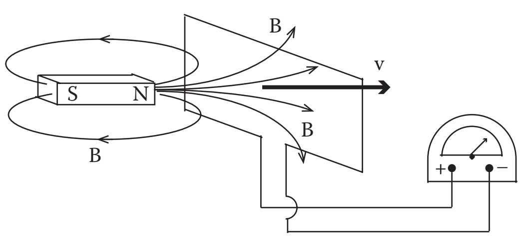


FIGURE 17.1 A magnet in motion induces an emf in a wire loop.

Experiment 2: Consider a circuit positioned close to a conducting loop (loop 2) which is attached to a galvanometer (Figure 17.2). When switch S is closed, a current I is quickly established in the circuit of loop 1. This current creates a magnetic field whose field lines "cut" through loop 2. An emf is induced in loop 2 while the current and its associated magnetic field are being established in loops 1 and 2 and the galvanometer deflects to the right of its "zero" position. When the current and its magnetic field achieve a steady state, the emf induced in loop 2 becomes zero. As long as the current in the circuit is in a steady state, no emf is induced in loop 2. If now switch S is opened, the current and its associated magnetic field quickly fall to zero. While the current is decreasing, an emf is induced in loop 2 and the galvanometer needle deflects to the left of its zero position. That

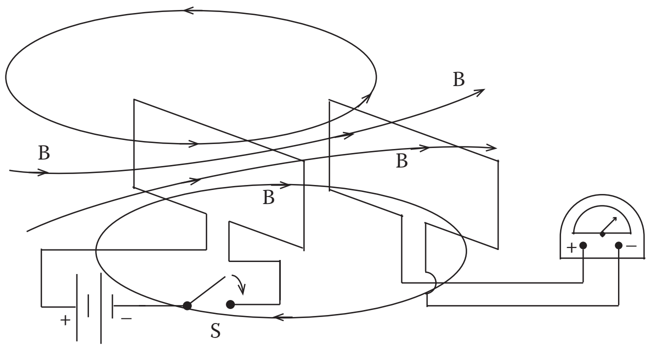


FIGURE 17.2 A circuit whose current is changing in time induces an emf in a close-proximity, closed conducting loop.

is, while the current is decreasing the polarity of the emf in loop 2 is opposite to its polarity when the current is increasing. When the current achieves a steady-state zero value, no emf is induced in loop 2. If the switch is now flipped "on" and "off" regularly, loop 2 has an emf induced with a given polarity when S is closed and the opposite polarity when S is opened.

Experiments 1 and 2 and other similar experiments led Faraday to conclude that an emf is induced in a loop, and indeed a region of space, whenever it experiences a changing magnetic field, or if the field is steady, a changing effective area. He astutely recognized that a time-varying quantity called the *magnetic flux* was the cause of this induced emf.

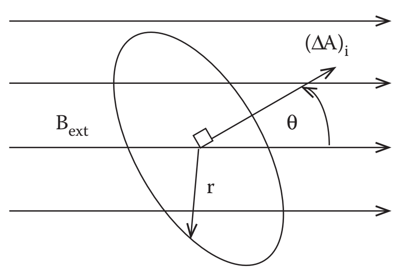
An increment of magnetic flux $\Delta\Phi$ is defined, in a fashion similar to the electric flux, as

$$(\Delta\Phi)_i = B_i(\Delta A)_i \cos\theta, \tag{17.1}$$

where B_i is the value of the magnetic field at the location of the incremental element of area $(\Delta A)_i$. Here, θ is the angle between the direction of B_i and the incremental element $(\Delta A)_i$. The flux through a finite area A is acquired by summing over all incremental elements $(\Delta A)_i$ that make up A. If the increments $(\Delta A)_i$ are infinitesimal, the sum in Equation 17.1 is replaced with an integral. Note that it is common jargon to say the flux "cuts" a given area. This just means that the lines of **B** pass through that area.

EXAMPLE 17.1

A circular conductor of radius $r = 6.00$ cm is placed in a uniform magnetic field of magnitude 1.00×10^{-2} T. The area of the plane of the loop makes an angle of 25.0° with the direction of **B** (the figure below). Determine the magnetic flux through the loop.



The magnetic flux through a "tilted" loop.

Solution

Summing ΔΦ in Equation 17.1 gives

$$\Phi = \sum B_i (\Delta A)_i \cos\theta = B\cos\theta \sum (\Delta A)_i. \tag{17.2}$$

Since B and θ are constant, they can be factored out of the sum. The sum is over the increments $(\Delta A)_i$. Performing the sum over all elements just yields the area of the loop, that is,

$$\sum (\Delta A)_i = A_{loop} = \pi r^2$$

Thus,

$$\Phi = B(A)_{loop}\cos\theta = (1.00\times10^{-2}\ T)\pi(0.0600\ m^2)^2\cos 25.0°$$

or

$$\Phi = 1.02\times10^{-4}\ Tm^2.$$

Note:

$$\text{If } \theta = 0.00°,\ \cos\theta = 1.00 \quad \text{and} \quad \Phi = 1.13\times10^{-4}\ Tm^2.$$

$$\text{If } \theta = 90.0°,\ \cos 90.0° = 0 \quad \text{and} \quad \Phi = 0.00\ Tm^2.$$

17.2 FARADAY'S LAW

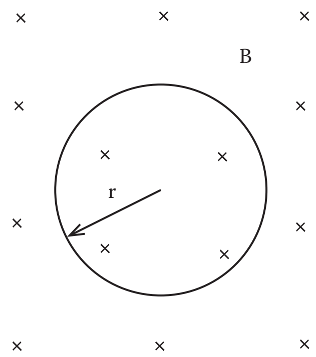
Faraday concluded that the induced emf ξ obeys the equation, named in his honor, and known as *Faraday's law*. That is,

$$\xi = -N\left(\frac{\Delta\Phi}{\Delta t}\right). \tag{17.3}$$

Here, N is the number of turns of a conducting loop that experiences a flux change (ΔΦ/Δt). Note that Faraday's law not only applies to conducting circuits but it also holds for empty regions of space (N = 1), in which no circuit is located (see Example 17.3). The negative sign in Equation 17.2 is used to determine the "polarity," that is, the direction of the concomitant conventional charge flow, of the emf and is known as *Lenz's law*. It is attributed to the Russian physicist Heinrich Lenz (1804–1865). Lenz's law implies that "the induced emf has a polarity that always opposes the flux change that caused it." This law is just a compliance with the conservation of energy and is verified experimentally.

EXAMPLE 17.2

A spatially uniform B field is increasing into the plane of the paper at the rate (ΔB/Δt) = 0.100 T/s. A conducting circular loop of radius r = 8.00 cm and resistance R = 0.900 Ω is placed in the field. Determine (a) the numerical value of the emf induced in the loop, (b) the induced current in the loop, and (c) the current direction clockwise or counterclockwise (the following figure).



The geometry for Example 17.2.

Solution

a. Faraday's law gives

$$\xi = -N\left(\frac{\Delta\Phi}{\Delta t}\right). \tag{17.4}$$

The $(\Delta A)_i$ are perpendicular to the plane of the loop and thus parallel to B_i. Thus $\theta = 0.00°$, $\cos\theta = 1.00$. B is spatially uniform and may be factored out of the sum. Equation 17.4 becomes

$$\xi = -(1)\frac{\Delta\left\{B\sum_i (\Delta A)_i\right\}}{\Delta t} = (-1)A\left(\frac{\Delta B}{\Delta t}\right) = (-1)\pi r^2\left(\frac{\Delta B}{\Delta t}\right) = -\pi(0.0800\,\mathrm{m})^2(0.100\,\mathrm{T/s}),$$

$$\xi = -2.01 \times 10^{-3}\ \mathrm{V} = -2.01\ \mathrm{mV}.$$

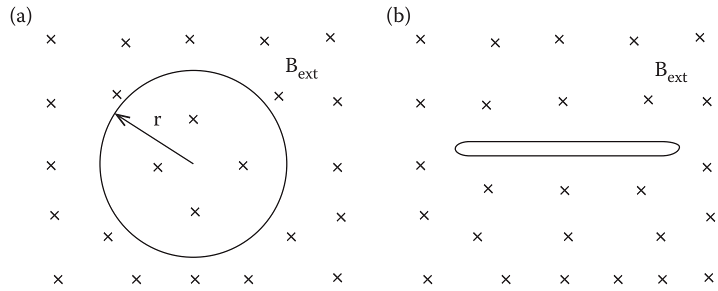
b. $\xi = IR$, so,

$$I = \frac{\xi}{R} = \frac{2.01\times 10^{-3}\ \mathrm{V}}{0.900\,\Omega} = 2.33\times 10^{-3}\mathrm{A} = 2.33\,\mathrm{mA}.$$

c. Here, Lenz's law implies that the induced emf, and its associated current, opposes the change that caused it. Remember that the induced current (I_i) gives rise to its own magnetic field B_{ind}. If the induced current is clockwise, the right hand rule-II gives B_{ind}, in the interior of the loop, into the paper. In this case, B_{ind} would add to B_{ext} and not oppose it. If I_{ind} is counterclockwise, its B_{ind} is directed out of the paper and opposes the increase in B_{ext}. So I_{ind} is counterclockwise.

EXAMPLE 17.3

A flexible circular conducting wire loop of radius r = 8.00 cm is placed in a spatially uniform magnetic field of value B = –0.150 T (the following figure, part (a)). The loop is clasped at opposite ends of a diameter and rapidly pulled such that its circumscribed area is reduced to 2.51×10^{-3} m² in 0.500 s (part (b)). Determine the value of the average emf induced in the loop during the 0.500 s time interval.



(a) A circular conductor in a uniform magnetic field. (b) An emf is induced in the loop during the change in its area.

Solution

$$\xi_{ind} = -N\left(\frac{\Delta\Phi}{\Delta t}\right) = -N\frac{\Delta\left\{\sum B_i(\Delta A)_i\cos\theta\right\}}{\Delta t}. \tag{17.5}$$

B is directed into the paper, the $(\Delta A)_i$ are in the plane of the loop and thus also directed into the paper, so $\theta = 0.00°$. Summing over the $(\Delta A)_i$s gives A, the area of the loop. Equation 17.5 can now be written as

$$\xi = -NB\left(\frac{\Delta A}{\Delta t}\right) = -NB\left(\frac{A_{final} - A_{initial}}{\Delta t}\right)$$

or

$$\xi_{ind} = (-1)(0.150\,\mathrm{T})\left(\frac{2.51\times 10^{-3}\,\mathrm{m}^2 - \pi(0.080\,\mathrm{m}^2)}{(0.500)\,\mathrm{s}}\right)$$

$$\xi_{ind} = -(0.150\,\mathrm{T})\left(\frac{(2.51 - 20.1)\times 10^{-3}\,\mathrm{m}^2}{(0.500)\,\mathrm{s}}\right) = 5.28\times 10^{-3}\,\mathrm{V} = 5.28\,\mathrm{mV}.$$

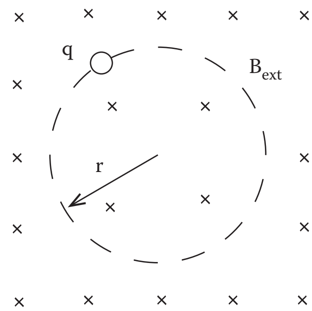
Note that since B is constant and the area was decreased, the flux "cutting" the loop was also decreased. Lenz's law implies that ξ_{ind} will generate an I_{ind} to oppose the flux decrease, that is, to add an increase to the flux. Thus, I_{ind} in the elongated loop (the figure above, part (b)) must be directed clockwise.

EXAMPLE 17.4

A spatially uniform **B** field is increasing into the paper at the rate $(\Delta B/\Delta t) = 0.100$ T/s. A charge of $q = 3.00\ \mu C$ and mass $= 3.13 \times 10^{-14}$ kg is placed in the field at a distance of $r = 8.00$ cm from its center of symmetry (the following figure). (a) Discuss the motion of the charge and (b) determine its velocity.

Solution

This example is somewhat similar to Example 17.2 except for the existence of a circular conductor in Example 17.2. Here, there is no conductor in the B field. An emf was induced around that conducting path and similarly an emf is induced around the circular path in this example. Indeed, an emf would be induced around any closed path through which the magnetic flux is changing. This illustrates the generality, power and beauty of Faraday's law. *Faraday's law applies to the space in which the flux is changing.*



Charges in a changing magnetic field experience an emf regardless of whether they are point charges or located in a conductor.

a. From Example 17.2, the emf induced around the path is

$$\xi = -A\left(\frac{\Delta B}{\Delta t}\right)$$

and it induces a current in the conductor. Positive charges flow from high to low potential, so a potential difference exists around the circular path. Thus, the 3.00-μC charge will move in a circle of radius 8.00 cm and in a counterclockwise sense.

b. The voltage difference around the path (V) must be equal to the induced emf ξ. Additionally, an electric field (**E**) must exist around the loop. **E** is related to V by the equation $|E| = V/\ell$ (Equation 14.11). Here ℓ is the circumference of the circular path. So,

$$\xi = V = |E|\ell = -A\left(\frac{\Delta B}{\Delta t}\right). \tag{17.6}$$

Solving for $|E|$ gives

$$|E| = \frac{A}{\ell}\left(\frac{\Delta B}{\Delta t}\right) = \frac{\pi r^2}{2\pi r}\left(\frac{\Delta B}{\Delta t}\right) = \frac{r}{2}\left(\frac{\Delta B}{\Delta t}\right). \tag{17.7}$$

To find the velocity, recognize that since the charge moves in a circular path, it is acted on by a centripetal force. Thus,

$$F = |E|q = m\frac{v^2}{r}$$

or

$$v = \left[\frac{rE\,q}{m}\right]^{1/2} = \left[\frac{q\,r^2}{2m}\left(\frac{\Delta B}{\Delta t}\right)\right]^{1/2}, \tag{17.8}$$

so

$$v = \left[\frac{(3.00 \times 10^{-6}\ \text{C})(0.080\ \text{m})^2(0.100\ \text{T/s})}{2(3.13 \times 10^{-14}\ \text{kg})}\right]^{1/2}$$
$$= 3.07 \times 10^4\ \text{m/s} = 0.19\ \text{c}.$$

EXAMPLE 17.5

A conducting metal bar is pushed along parallel conducting rails with a velocity v. At one end, the ends of the rails are electrically connected to a galvanometer such that the bar, rails, and galvanometer form a completed circuit. This circuit is immersed in a spatially uniform external field B_{ext} (the figure below). Derive an expression, in terms of the given parameters, for the emf induced around the loop.

Solution

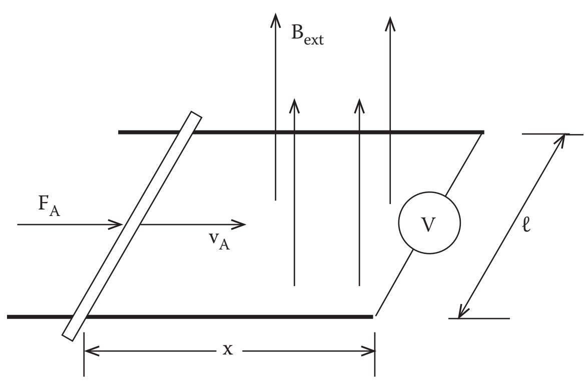
Since B_{ext} and (ΔA) interior to the loop are both directed upward, $\theta = 0.00°$. Here, Faraday's law can be written as

$$\xi_{ind} = -N\left(\frac{\Delta\Phi}{\Delta t}\right) = -N\frac{\Delta(B_{ext}A)}{\Delta t} = -NB_{ext}\left(\frac{\Delta A}{\Delta t}\right).$$

At any instant, $A = \ell x$, so

$$\xi_{ind} = -NB_{ext}\frac{\Delta(\ell x)}{\Delta t} = -NB_{ext}\ell\left(\frac{\Delta x}{\Delta t}\right) = -NB_{ext}\ell v_A.$$

Here $N = 1$. Since the flux through the loop is decreasing as the bar moves toward the galvanometer, the ξ_{ind} around the loop will create an I_{ind} to *add* to the decreasing flux. Thus, I_{ind} is directed from top to bottom in the bar or counterclockwise, from above, in the loop.



A conducting rod being pushed through a magnetic field with an applied velocity v_A.

17.3 ELECTRIC GENERATORS

The operation of an electric generator can be understood on the basis of Faraday's law. Consider a rectangular conducting loop (Figure 17.3a).

The loop, called an *armature*, is placed in a spatially uniform magnetic field and mounted with provision to allow it to rotate about a centered axis (Figure 17.3b). The area circumscribed by the loop (A_o) and the field (B) are constant, but the magnetic flux through the loop changes as it is rotated.

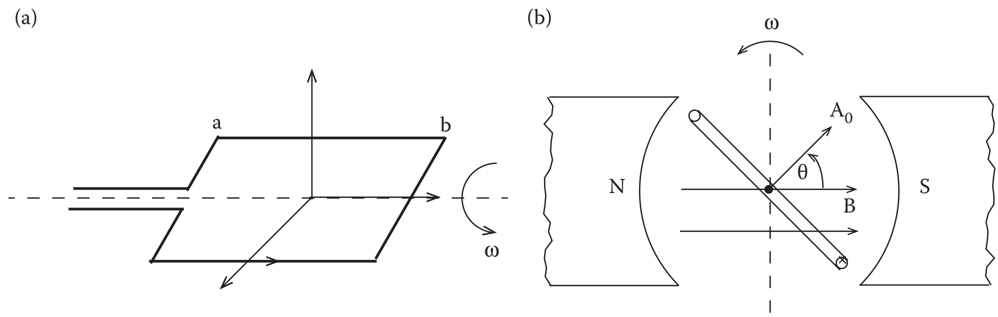


FIGURE 17.3 (a) The geometry for application of Faraday's law. (b) A conducting loop free to rotate in a magnetic field.

The changing flux is due to the changing projection of A_o along B. Faraday's law gives

$$\xi_{ind} = -N\left(\frac{\Delta\Phi}{\Delta t}\right) = -\frac{\Delta\left\{\sum B_i(\Delta A)_i \cos\theta_i\right\}}{\Delta t}. \tag{17.9}$$

The sum over $(\Delta A)_i$ yields A_o, so Equation 17.9 can be written as

$$\xi_{ind} = -NBA_o \frac{\Delta(\cos\omega t)}{\Delta t} \tag{17.10}$$

where $\theta = \omega t$. The time rate of change of cos ωt, in the limit that (Δt) becomes infinitesimally small, is evaluated via calculus. The result is (ω sin ωt). The emf around the loop of the generator is thus,

$$\xi_{ind} = -NBA_o\omega \sin\omega t. \tag{17.11}$$

Notice that ξ_{ind} depends on N and ω. Constructing the armature of many turns and rotating it more rapidly produces a larger ξ_{ind}. If the generator is connected to an external circuit, it becomes a source of charge flow to the circuit. The method by which the armature is connected determines whether the charge flow will be alternating current (ac) or direct current (dc).

17.3.1 AC Generator

If the ends of the simple armature are rigidly connected to conducting rings (Figure 17.4) that rotate with it, the polarity of the voltage across the resistor and the current through it will alternate in direction. To see this, note that when the loop segment (ab) in Figure 17.3a is situated in the left half-plane, the induced current is directed out of the paper, that is, from point b to point a. When segment

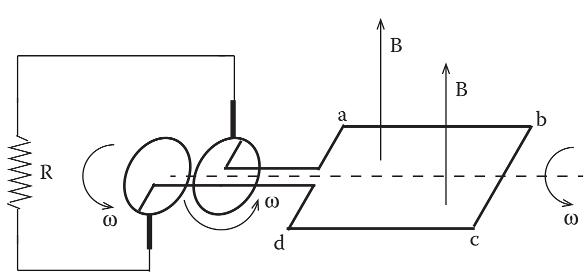


FIGURE 17.4 The "slip-ring" connection to the armature results in an ac voltage output.

(ab) is in the right half-plane, the current is directed into the paper, that is, from point a to point b. So when segment (ab) crosses from one half-plane to the other, the current in that segment reverses direction. Since that armature segment is connected to the simultaneously rotating ring, the current in the resistor will reverse direction every one-half rotation, that is, every one-half cycle. The same discussion applies to segment (cd). Since ξ_{ind} varies as sin ωt, I_{ind} in the armature and "fed" to the resistor also varies as sin ωt (Figure 17.5).

The circuit symbol for an ac generator is sketched on the right side of the figure.

17.3.2 DC Generator

If the ends of the simple armature are rigidly connected to conducting semicircles (half rings or split-rings) that rotate with it (Figure 17.6), the current through the external resistor will be pulsating direct current, that is, dc (Figure 17.7). The half rings or split-rings are called a *commutator*. They rotate with the armature and slide over contacts are called *brushes*. This arrangement produces pulsating dc. Whichever segment of the armature, (ab) or (cd), is in the left half-plane in Figure 17.3b will have an induced current directed out of the paper. That segment, via the split-ring commutator, is always connected to the brush on the left. Thus, the current through the resistor is always from positive to negative. It does not reverse direction every half cycle as does the current from the ac generator.

A practical generator has an armature that consists of many turns of wire, not the simple one turn as shown in Figures 17.3, 17.4 and 17.6. The turns are wound on an iron core that directs and vastly increases the magnetic flux.

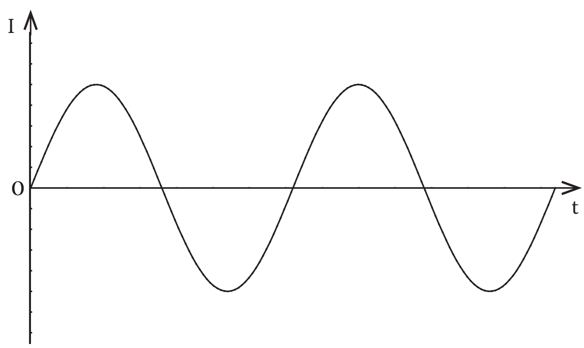


FIGURE 17.5 The sinusoidal output of an ac generator.

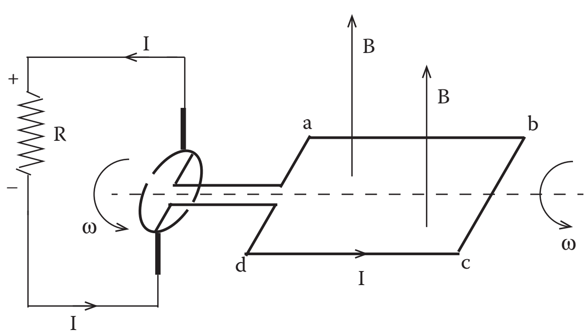


FIGURE 17.6 The "commutator" connection to the armature results in a pulsating dc voltage output.

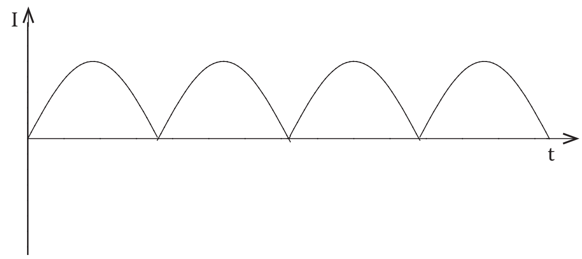


FIGURE 17.7 The pulsating output of a dc generator.

EXAMPLE 17.6

It is desired to design a single plane, N-turn ac generator, typical of your automobiles alternator, with the following parameters: armature length = 7.00 cm, armature width = 8.00 cm, B = 0.120 T, $\omega = 2000$, rpm = 209 rad/s. For what value of N will the induced emf have an amplitude of 14.2 V?

Solution

Using Equation 17.11 with $\sin \omega t = 1.00$, $\xi_{ind} = \xi_o$, and solving for N gives

$$N = \frac{\xi_o}{BA_o\omega} = \frac{(14.2\ \text{V})}{(0.120\ \text{T})(0.0700\ \text{m})(0.0800\ \text{m})(209\ \text{rad/s})},$$

$$N = 101\ \text{turns}.$$

17.4 SELF-INDUCTANCE

Faraday's law implies that an emf is induced in that region of space where the magnetic flux changes with time. When a circuit element such as an ideal solenoid is connected to an ac source, as in Figure 17.8, it is immersed in its own alternating magnetic field and experiences a changing magnetic flux. To see this, note that an ideal solenoid has B confined to its interior. For a solenoid of length ℓ with N turns

$$B = \frac{\mu_o NI}{\ell}. \tag{17.12}$$

Since the current varies with time, so does B and therefore Φ.

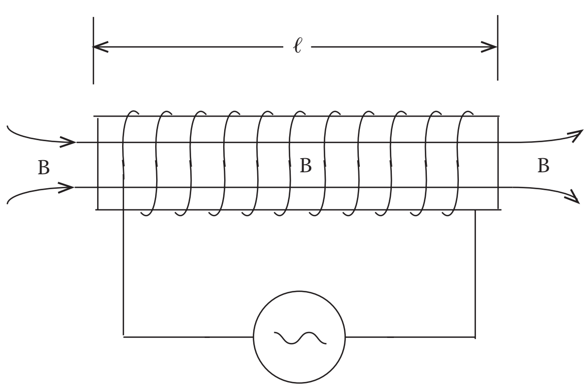


FIGURE 17.8 An ideal solenoid connected to an ac source.

Let ξ_A equals the externally applied emf from the generator and ξ_i is the emf induced in the solenoid. If the resistance of the solenoid is assumed negligible, then applying Kirchhoff's loop rule gives

$$\sum \xi = 0 = \xi_A - \xi_i$$

or

$$\xi_A = \xi_i = -N\left(\frac{\Delta\Phi}{\Delta t}\right) = -N\frac{\Delta(BA)}{\Delta t} = -NA\left(\frac{\Delta B}{\Delta t}\right). \tag{17.13}$$

Here, A is the area of the interior cross section. Note that ξ_i opposes the change that causes it, so it opposes the source voltage ξ_A. For this reason, ξ_i is called a "back emf." Inserting Equation 17.12 into Equation 17.13 gives

$$\xi_i = -NA\frac{\Delta}{\Delta t}\left[\frac{\mu_o NI}{\ell}\right] = -\left(\frac{\mu_o N^2 A}{\ell}\right)\left(\frac{\Delta I}{\Delta t}\right). \tag{17.14}$$

The terms in (N^2A/ℓ) are geometrical. Equation 17.14 is usually written as

$$\xi_i = -L\left(\frac{\Delta I}{\Delta t}\right), \tag{17.15}$$

where, in this special case of the ideal solenoid,

$$L = \left(\frac{\mu_o N^2 A}{\ell}\right). \tag{17.16}$$

The quantity L is called the *self-inductance*, or in short, the inductance. From Equation 17.15

$$L = \left|\frac{-\xi_i}{(\Delta I/\Delta t)}\right|. \tag{17.17}$$

So, L is a measure of how much back emf is induced per a given $(\Delta I/\Delta t)$. If $\xi_i \equiv 1.00$ V when the current changes at a rate $(\Delta I/\Delta t) = 1.00$A/s, then L = 1.00 Henrys. The expression for L, for geometries other than solenoids, takes a different form than Equation 17.16. When placed in an ac circuit, these elements are called *inductors*. Note that ξ_i opposes ξ_A.

EXAMPLE 17.7

A small solenoid has an inductance of L = 0.670 mH. Determine its back emf ξ_i when the current through it changes at a rate of 8.00 A/s.

Solution

Using Equation 17.15

$$\xi_i = -L\left(\frac{\Delta I}{\Delta t}\right) = -(0.670\times10^{-3}\ \text{H})(8.00\ \text{A/s}) = -5.36\times10^{-3} = -5.36\ \text{mV}.$$

17.5 TRANSFORMERS

A transformer is a device that exploits Faraday's law for several purposes, one of which is to increase (step-up) or decrease (step-down) applied voltages. Figure 17.9 shows a fundamental transformer. A ferromagnetic core, typically iron or AlNiCo, is shaped as a rectangle with its center material removed. Insulated wire is wound around one segment of the core, with N_1 turns, and connected to an ac source. This coil is called the *primary coil*. A second coil of insulated wire, called the *secondary coil*, and consisting of N_2 turns is wound around the core and connected to a resistive "load" R_L. The ac source produces a time-varying magnetic flux in the primary coil. This alternating flux, which is enhanced by and mostly confined to the core, also "cuts" the secondary coil. Note that the ac source and the primary coil are a closed loop.

Assuming negligible resistance in this primary loop and applying Kirchhoff's loop rule to it gives

$$\sum(\text{potential changes}) = 0 = \xi_A - \xi_1$$

or

$$\xi_A = \xi_1 = -N_1\left(\frac{\Delta\Phi}{\Delta t}\right)_1. \tag{17.18}$$

The core transmits the changing flux to the secondary coil and induces an emf, ξ_2 in it. Thus,

$$\xi_2 = -N_2\left(\frac{\Delta\Phi}{\Delta t}\right)_2. \tag{17.19}$$

In an ideal transformer, all the magnetic flux is confined to the core so

$$\left(\frac{\Delta\Phi}{\Delta t}\right)_2 = \left(\frac{\Delta\Phi}{\Delta t}\right)_1. \tag{17.20}$$

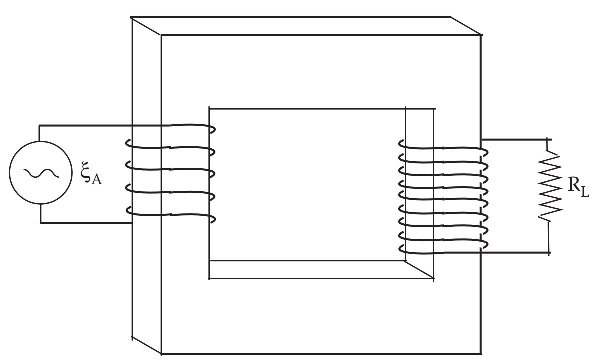


FIGURE 17.9 A transformer can increase (step-up) or decrease (step-down) an ac voltage applied to its primary coil.

Using Equations 17.18 and 17.19 in Equation 17.20 gives

$$\frac{\xi_1}{N_1} = \frac{\xi_2}{N_2} \quad \text{or} \quad \xi_2 = \left(\frac{N_2}{N_1}\right)\xi_1. \tag{17.21}$$

If $N_2 > N_1$, Equation 17.21 gives $\xi_2 > \xi_1$ and the transformer is called a *step-up transformer*. The voltage output of the secondary coil is greater than the voltage input to the primary. When $N_2 < N_1$, $\xi_2 < \xi_1$, that is, the secondary voltage is less than the primary voltage, the transformer is called a *step-down transformer*.

If power losses in the transformer are negligible, the power output of the secondary (P_2) is equal to the power input to the primary (P_1). Thus,

$$P_2 = P_1 \quad \text{or} \quad I_2\xi_2 = I_1\xi_1, \tag{17.22}$$

then

$$I_2 = I_1\left(\frac{\xi_1}{\xi_2}\right) = I_1\left(\frac{N_1}{N_2}\right). \tag{17.23}$$

Equation 17.23 shows that although $\xi_2 > \xi_1$ in a step-up transformer, $I_2 < I_1$. Thus, a step-up transformer steps up the secondary voltage but steps down the secondary current.

EXAMPLE 17.8

It is desired to design a transformer with an output voltage of $\xi_2 = 15{,}000$ V (needed to "fire" the spark plugs of an automobile engine) when the input is $\xi_1 = 12.0$ V. The number of turns in the primary is $N_1 = 24.0$. Determine N_2.

Solution

Using Equation 17.21

$$N_2 = N_1\left(\frac{\xi_2}{\xi_1}\right) = (24.0)\left(\frac{15{,}000\text{ V}}{12.0\text{ V}}\right) = 2500\text{ turns.}$$

PROBLEMS

17.1 A transformer connected to a 120.0-V ac line is supplying 6.00 V to a radio. The number of turns in the primary is 300.0. Determine the number of turns in the secondary.

17.2 A rectangular-shaped conducting coil of 350.0 turns has a length of 25.0 cm and a width of 20.0 cm. The coil is immersed in a spatially uniform magnetic field that increases in value from 0.150 T to 0.550 T in 0.700 s. The plane of the coil is perpendicular to the field direction. Determine the magnitude of the emf induced in the coil.

17.3 An inductor has a self-inductance of 0.800 Hs. The current through it is changing at the rate of 35.0 A/s. Determine the magnitude of the emf induced in the coil.

17.4 A circular loop of wire with radius r = 3.00 cm is placed in a spatially uniform magnetic field of magnitude B = 0.120 T. Determine the flux through the loop for the following orientations:

a. The plane of the loop perpendicular to the direction of the magnetic field
b. The plane of the loop parallel to **B**
c. The plane of the loop makes an angle of 45.0° with the direction of **B**

17.5 A rectangular-shaped conducting coil, with 35.0 turns, has a length of 25.0 cm and a width of 20.0 cm. The coil is immersed in a 0.550 T, spatially uniform magnetic field. The coil is clasped at two diagonal corners that are pulled oppositely outward so that the "elongated" coil encloses *zero* area. This action occurs in 0.200 s. Determine the magnitude of the emf induced in the coil during elongation.

17.6 A 248-turn rectangular loop (armature) encloses an area of 0.400 m^2. The coil is rotating at an angular speed of 15.0 rad/s about an axis that is perpendicular to a magnetic field of 0.112 T. Determine the maximum emf induced in this generator.

17.7 It is determined that an inductor has an emf of 108 V induced in it when the current through it is changing at the rate of 240 A/s. Determine the value of its self-inductance.

17.8 A doorbell requires 10.0 V to operate. This voltage is supplied by a step-down transformer connected to standard 120.0 V house voltage. The transformer primary consists of 200.0 turns. How many turns should the secondary have?

17.9 It is desired to design a 60.0-Hz ac generator whose maximum voltage output is 440.0 V. The armature should have 80.0 similar turns, each circumscribing an area of 0.127 m^2. Calculate the magnitude of the magnetic field in which the armature should rotate.

17.10 An emf of 8.00 V is induced across an ideal solenoid when the current through it is changing at a rate of 25.0 A/s. Determine the solenoid's self-inductance.

17.11 It is desired to design an ideal solenoid that is 25.0 cm long, whose identical turns circumscribe an area 1.25 cm^2 that has a self-inductance of $L = 3.20 \times 10^{-5}$ Henries. How many turns should it have?

17.12 A circular coil of radius 8.00 cm is wound with 80.0 identical turns. The coil is placed, with its plane perpendicular to a time-varying, spatially uniform magnetic field. At what rate must the field change so that an emf of 1.50 V is induced in the coil.

17.13 An air-core solenoid has a self-inductance of $L = 2.80$ mH. If the current through it is changing at the rate of 0.500 A/s, what is the value of the emf induced in the solenoid?

17.14 An armature has 90.0 identical turns, each circumscribing an area of 1.13×10^{-2} m^2. The armature is rotated in a magnetic field of average value 0.130 T. At what angular velocity should it be rotated to produce a maximum output of 12.0 V?

17.15 The alternator in your automobile "puts out" an approximate 13.8 V when rotated at 1000 rev/min. The 80.0 armature turns each circumscribe an area approximately 5.50×10^{-3} m^2. Determine the magnitude of the magnetic field in which the armature is rotating.

18 Alternating Current Electric Circuits

18.1 INTRODUCTION

The invention and development of the alternating current (ac) generator helped lead to the electrification of the United States in the early twentieth century. The advantages of sinusoidal ac signals over direct current (dc) and other alternating electrical signals are addressed in this chapter.

To review, recall that dc or ac refers to electric current directions or to source voltage polarity. Figure 18.1 shows a simple dc circuit. The positive terminal (polarity) of the battery is always connected to the same side of the resistor. When switch S is closed, conventional charge will leave the battery and flow through the resistor R from left to right.

Since positive charge flows from high to low potential, the polarity of the resistor is left-side positive and right-side negative as shown in Figure 18.1. A plot of both the voltage drop across and the current through the resistor is shown in Figure 18.2. Note that the polarity of the source and the voltage drop across R remain constant in time.

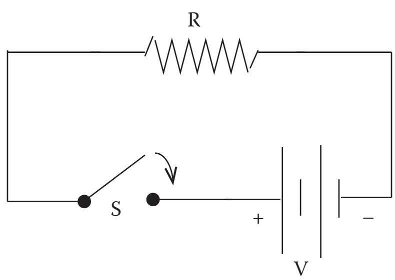


FIGURE 18.1 A simple dc circuit. Charge flows in only one direction.

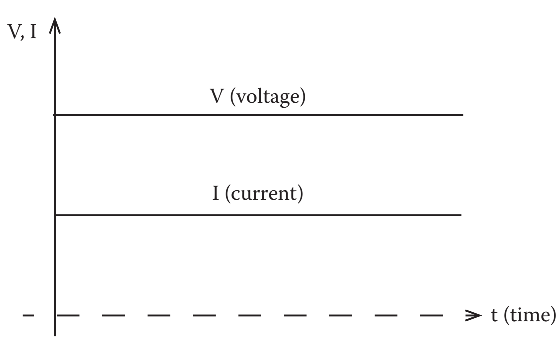


FIGURE 18.2 Current direction and voltage polarity are constant in time in a dc circuit.

18.2 ALTERNATING SIGNALS

An ac circuit is one whose source voltage polarity alternates in time between positive and negative. There exist various forms of alternating signals. A few simple ac circuits are shown in Figure 18.3. The sinusoidal signal can be acquired easily from an ac generator of the type discussed in Chapter 17 and is the form used by electrical generating plants to distribute electric power to the public. Why ac? Alternating signals allow for less costly distribution of electric power than does dc, because the step-up property of transformers can be exploited to minimize dissipative heat losses. The following simplified example will illustrate this concept.

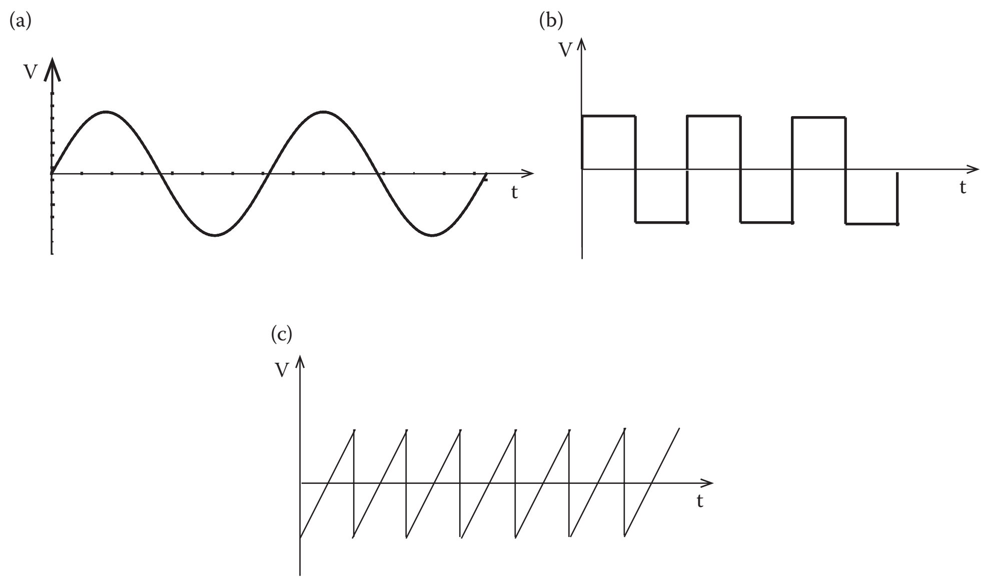


FIGURE 18.3 Some time-varying voltage forms in an ac circuit: (a) sinusoidal, (b) square, and (c) sawtooth.

EXAMPLE 18.1

A toaster oven in a home that is 30.0 miles from an electric power generating station has a resistance $R = 15.0\ \Omega$ and requires 8.00 A to operate at 120.0 V. The resistance of each of the two transmission lines stretching from the station to the home is $4.80\ \Omega$.

a. If the generator at the power station supplied 8.00 A at 120.0 V (in reality, the station generator voltage is approximately 100 times this value), how much power would the toaster receive?
b. If now a step-up transformer with $N_2 = 1000N_1$ was installed at the station and a step-down transformer with $N_1 = 1000N_2$ was installed near the home, how much power would the toaster receive?

Solution

a. The power output of the generator, required for the toaster, is

$$P_G = I_G V_G = (8.00\ \text{A})(120.0\ \text{V}) = 960\ \text{W}.$$

The power lost to heating the two wires that deliver power to the home, 30.0 miles from the station, is

$$P_{lost} = I^2R = (8.00\ \text{A})^2\ [2(4.80\ \Omega)] = 614\ \text{W}.$$

The power available to the toaster is

$$P_{avail} = P_G - P_{lost} = 960\ \text{W} - 614\ \text{W} = 346\ \text{W}.$$

b. If the 120.0 V at the power station is fed into the primary coil of the transformer, the output of the secondary (V_2) will be

$$\xi_2 = \left(\frac{N_2}{N_1}\right)\xi_1 = (1000)(120.0\ \text{V}) = 1.20 \times 10^5\ \text{V}.$$

The transformer steps up the voltage, but correspondingly steps down the current output. The current drawn is

$$I = \frac{P_G}{V_G} = \frac{960\ \text{W}}{1.20 \times 10^5\ \text{V}} = 8.00 \times 10^{-3}\ \text{A} = 8.00\ \text{mA}.$$

Now, the power lost to heating the transmission lines is

$$P_{lost} = I^2R = (8.00 \times 10^{-3}\ \text{A})^2[2(4.80\ \Omega)] = 6.14 \times 10^{-4}\ \text{W}.$$

The power available to the toaster is

$$P_{avail} = P_G - P_{lost} = 960\ \text{W} - 0.614 \times 10^{-3}\ \text{W} \approx 960\ \text{W}.$$

Negligible power is lost in transmission, of even large amounts of power, if the current can be kept small. Near the home, the voltage can be stepped down with a corresponding step up in current. The values from the step-down transformer are

$$V_2 = \xi_2 = \left(\frac{N_2}{N_1}\right)V_1 = \left(\frac{1}{1000}\right)(1.20 \times 10^5\ \text{V}) = 120\ \text{V}$$

and

$$I_2 = I_1\left(\frac{N_1}{N_2}\right) = (8.00 \times 10^{-3}\ \text{A})\left(\frac{1000}{1}\right) = 8.00\ \text{A}.$$

So, the ac in the primary coil of a transformer creates an alternating magnetic flux in both the primary and secondary coils. Because of Faraday's law, this allows a current reduction in the secondary coil and thus minimizes conduction heat losses in transmission lines. With dc, there is no alternating magnetic flux so a transformer cannot function as a step-up, step-down device. The circuit symbol for a transformer is .

18.3 PHASE RELATIONS IN SIMPLE AC CIRCUITS

In this section, the phase relations between current through and voltage drop across three simple circuit elements are discussed. Additionally, only sinusoidal current and voltage signals are considered.

In Chapter 17, the output electromotive force (emf) of an ac generator was shown by Equation 17.11 as

$$\xi = NBA\omega \sin \omega t.$$

When connected to an external circuit, the generator supplies a time-varying potential difference that can be written as

$$\xi = V = V_0 \sin \omega t. \tag{18.1}$$

Here $V_0 = NBA\omega$ is called the *voltage amplitude* or just the amplitude for short and $\omega = 2\pi f$, where f is the frequency that specifies the number of complete polarity oscillations per second.

The SI unit of frequency is Hertz (Hz). The "time" in Equation 18.1 is measured with respect to some reference time and the reference time was chosen to be $t = 0$ when $V = 0$ (Figure 18.4a). Note that we are free to choose when we start our signal timing clock, so if the reference time had been chosen such that $V = V_0$ when $t = 0$, the generator voltage would be expressed as (Figure 18.4b).

$$V = V_0 \cos \omega t. \tag{18.2}$$

The cos ωt description leads (i.e., V_0 comes earlier in time) the sin ωt by one-fourth of a cycle. These two descriptions are "out-of-step" and we say there is a phase difference between them. Correspondingly two out-of-step signals with the same reference time are said to have a *phase difference* or be *out of phase*. *Note:* Since $\cos(A \pm B) = \cos A \cos B - \sin A \sin B$, Equation 18.2 can be written as

$$V = V_0 \sin\left(\omega t - \frac{\pi}{2}\right). \tag{18.3}$$

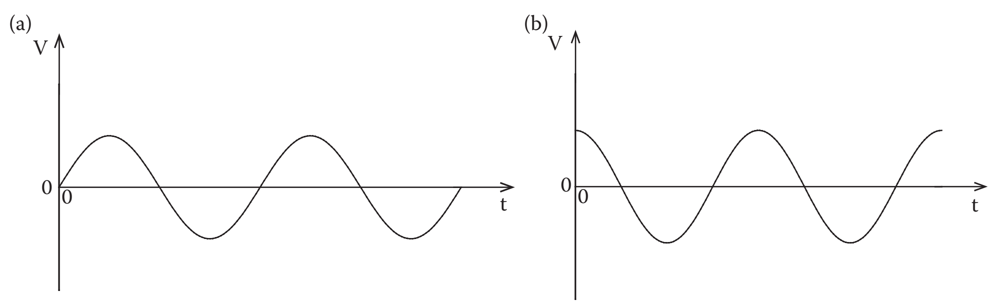


FIGURE 18.4 (a) If $V = 0$ when the time $t = 0$, then a sine function is a convenient description of the voltage signal. (b) If $V = V_0$ when $t = 0$, a cosine function is a convenient description.

18.3.1 Voltage–Current Phase Differences

18.3.1.1 Simple Resistive Circuit

Consider the simple ac circuit of Figure 18.5. With switch S closed, Kirchhoff's loop rule gives

$$V - V_R = V - IR = 0$$

or

$$V = IR = V_0 \sin \omega t. \tag{18.4}$$

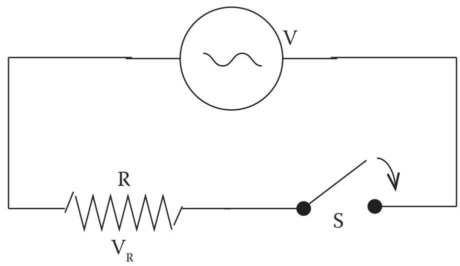


FIGURE 18.5 The simplest ac circuit.

Solving for the current gives

$$I = \left(\frac{V_0}{R}\right)\sin \omega t = I_0 \sin \omega t. \tag{18.5}$$

Since both I and V have the same reference time and vary as sin ωt, they are *in phase*. Figure 18.6 shows that both V(t) (curve a) and I(t) (curve b) reach their maximum amplitude at the same time. So, the voltage drop across the resistor and the current through the resistor are in phase with each other.

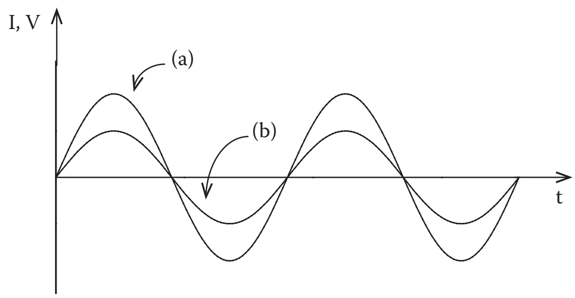


FIGURE 18.6 The current through and the voltage drop across a resistor are in phase with each other.

18.3.1.2 Simple Capacitive Circuit

Consider the simple ac circuit of Figure 18.7. With S closed, Kirchhoff's loop rule gives

$$V - V_C = V - \frac{q}{C} = 0. \tag{18.6}$$

Now,

$$V = V_0 \sin \omega t. \tag{18.7}$$

To determine I(t), rearrange Equation 18.6 to

$$q = CV = CV_0 \sin \omega t$$

and

$$I = \frac{\Delta q}{\Delta t} = CV_0 \frac{\Delta}{\Delta t}(\sin \omega t).$$

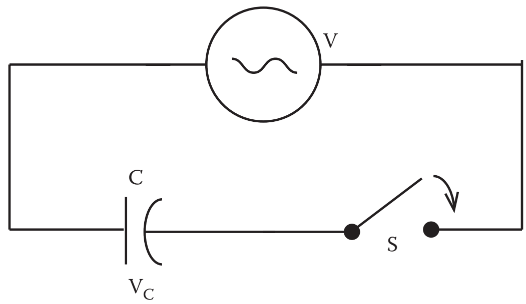


FIGURE 18.7 A simple capacitive ac circuit.

By the use of calculus,

$$\frac{\Delta}{\Delta t}(\sin \omega t) = \omega\cos \omega t.$$

So,

$$I = CV_0\omega \cos \omega t = I_0 \cos \omega t. \tag{18.8}$$

Comparing Equations 18.7 and 18.8, it is seen that the current "through" the capacitor (curve b) leads the voltage drop across it (curve a) by one-fourth cycle. The two quantities are "out of phase." We have stated that the current leads the voltage, but it is equivalent to state that the voltage "lags" the current. This is shown in Figure 18.8a and b.

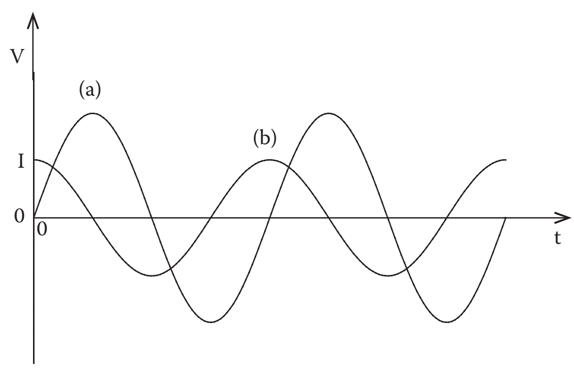


FIGURE 18.8 The ac voltage across a capacitor lags the current "through" it by one-fourth cycle.

18.3.1.3 Simple Inductive Circuit

Consider the simple ac circuit in Figure 18.9. With S closed, Kirchhoff's loop rule gives

$$V - V_L = V - L\left(\frac{\Delta I}{\Delta t}\right) = 0. \tag{18.9}$$

Rearranging

$$L\left(\frac{\Delta I}{\Delta t}\right) = V = V_0\sin \omega t. \tag{18.10}$$

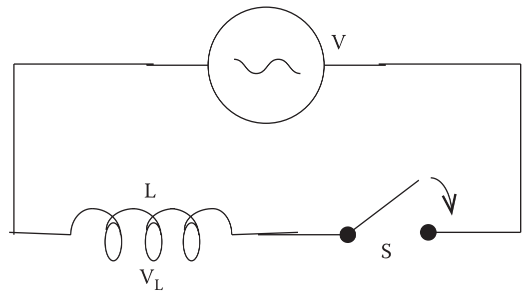


FIGURE 18.9 A simple inductive ac circuit.

So,

$$\left(\frac{\Delta I}{\Delta t}\right) = \left(\frac{V_0}{L}\right)\sin \omega t.$$

An expression for I(t) can be obtained from Equation 18.10 via calculus. The result is

$$I = -\left(\frac{V_0}{L\omega}\right)\cos \omega t = -I_0\cos \omega t. \tag{18.11}$$

Equations 18.10 and 18.11 are plotted in Figure 18.10. The graph shows that V leads I by one-fourth cycle.

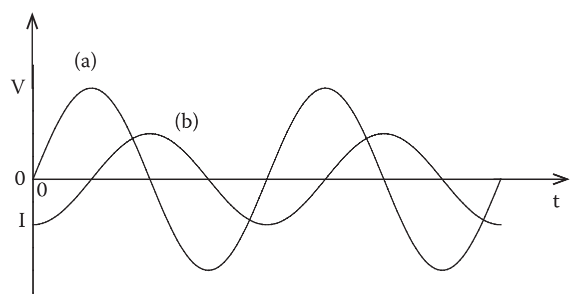


FIGURE 18.10 The ac voltage across an inductor leads the current through it by one-fourth cycle.

Analysis of the simplest ac circuits yielded the phase relations between voltage across and current through a resistor, capacitor, and inductor. In summary,

Resistor: V_R and I_R are in phase: $I_0 = \left(\frac{V_{0R}}{R}\right)$.

Capacitor: V_C lags I_C by 90°: $I_0 = \omega CV_{0C}$.

Inductor: V_L leads I_L by 90°: $I_0 = \left(\frac{V_{0L}}{\omega L}\right)$.

Here, the third column shows the relation between the *current amplitude* (I_0) and the voltage amplitude for each of the circuit elements. The equations may be rearranged to the form

$$V_{0R} = I_0R, \quad V_{0C} = \frac{I_0}{\omega C}, \quad V_{0L} = I_0\omega L. \tag{18.12}$$

It is convenient to write the last two equations of Equation 18.12 in the same form as the first, that is,

$$V_{0C} = I_0 X_C \text{ and } V_{0L} = I_0 X_L,$$

where

$$X_C \equiv \frac{1}{\omega C} \tag{18.13}$$

and

$$X_L \equiv \omega L. \tag{18.14}$$

X_C is called the *capacitive reactance* and X_L the *inductive reactance*.

18.4 RCL SERIES CIRCUIT

Consider the ac circuit in Figure 18.11. The resistor, capacitor, and inductor are connected in series with each other and with the source. If R, C, L, and V_S are known, the circuit current I_S and phase angle between I_S and V_S can be determined.

To analyze this problem, it is convenient to construct a so-called phasor diagram that helps us to account for these phase differences (see Figure 18.12). In Figure 18.12a, V_{0R} is drawn along the abscissa, similar to a vector, at time $t = 0$. Since V_{0L} leads V_{0R} by one-fourth cycle, that is, by 90°, it is drawn along the ordinate axis. V_{0C} lags V_{0R} by 90°, so it is drawn in opposition to V_{0L}. In Figure 18.12b, the vectors are drawn at an angle $\theta = \omega t$, that is, they are rotated about the origin. The projection of these vectors along the ordinate axis gives the time dependence of that quantity. So,

$$V_R = V_{0R} \sin \omega t,$$

$$V_L = V_{0L} \sin(\omega t + 90°),$$

$$V_C = V_{0C} \sin(\omega t + 270°).$$

To determine the circuit current I_S and the phase angle between V_S and I_S, a phasor diagram including V_{0S} should be constructed. Kirchhoff's loop rule Z gives the relation between the instantaneous values of the voltages. That is,

$$\sum V_{\text{loop}} = V_S - V_R - V_C - V_L = 0$$

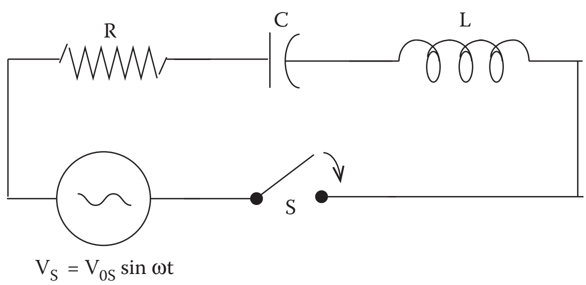


FIGURE 18.11 An RCL series, ac circuit.

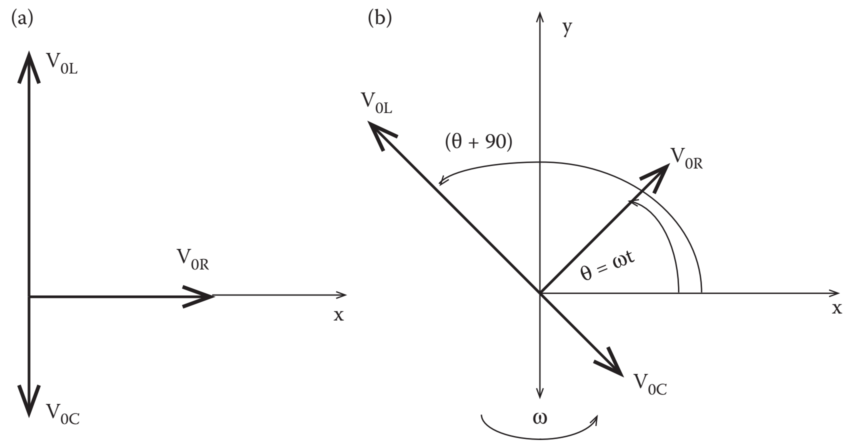


FIGURE 18.12 (a) A convenient way to represent the phase differences between V_{0R}, V_{0C}, and V_{0L}. (b) The projection of the amplitudes along the vertical axis gives the time dependence of V_R, V_C, and V_L.

or

$$V_S = V_R + V_C + V_L. \tag{18.15}$$

The sum on the right side of Equation 18.15 is acquired by *vectorally* adding the terms in Figure 18.12b. Note that V_{0L} leads V_{0C} by 180° or one-half cycle that results in a subtraction. So V_{0S} is the *vector* sum of the amplitudes of the other voltages. By the Pythagorean theorem

$$V_{0S}^2 = V_{0R}^2 + (V_{0L} - V_{0C})^2. \tag{18.16}$$

Using the third column of the equations under Equation 18.11 and Equations 18.13 and 18.14, Equation 18.16 can be written as

$$V_{0S}^2 = (I_0R)^2 + (I_0X_L - I_0X_C)^2$$

or

$$V_{0S} = I_0\sqrt{R^2 + (X_L - X_C)^2}. \tag{18.17}$$

The right side of Equation 18.17 can be put in the form of Ohm's law for a dc circuit by defining a quantity called the *circuit impedance* as

$$Z = \sqrt{R^2 + (X_L - X_C)^2}. \tag{18.18}$$

So, Equation 18.17 can be written as

$$V_{0S} = I_0Z \tag{18.19}$$

or

$$I_0 = \frac{V_{0S}}{Z}. \tag{18.20}$$

Equation 18.20 shows that I_0 is inversely proportional to Z. Thus, for a given source voltage V_{0S}, large Z implies relatively small I_0, that is, Z is a hindrance to charge flow just as resistance R is in dc circuits. Note that Z depends on source frequency ω through the terms X_C and X_L. The SI units of Z are ohms (Ω).

In Section 18.3, it was shown that the current through a resistor is in phase with the voltage drop across it. Thus, in Figure 18.13, the phase of I_S is equal to the phase of IR, but not equal to the phase of V_S. The phase difference between V_S and I_S, from Figure 18.13, is

$$\varphi = \tan^{-1}\frac{(V_{0L} - V_{0C})}{V_{0R}} = \tan^{-1}\frac{(I_0X_L - I_0X_C)}{I_0R} \tag{18.21}$$

or

$$\varphi = \tan^{-1}\frac{(X_L - X_C)}{R}. \tag{18.22}$$

Equations 18.21 and 18.22 show that the vector-like voltage amplitudes may be replaced by a similar triangle whose sides are reactances and whose hypotenuse is the circuit impedance (see Figure 18.14). In this section, equations for the circuit current I_S and phase angle φ were

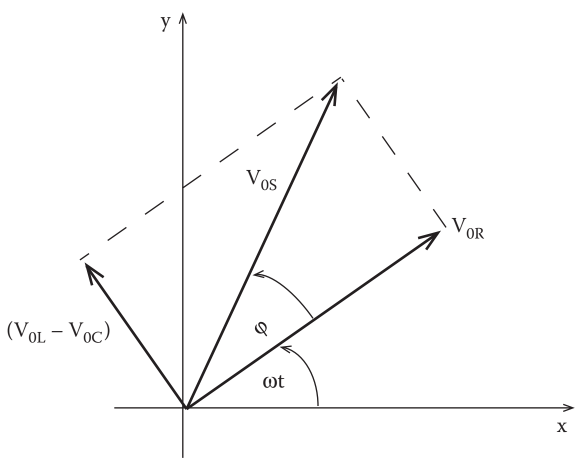


FIGURE 18.13 The diagram for determining angle φ, between the total circuit current and source voltage.

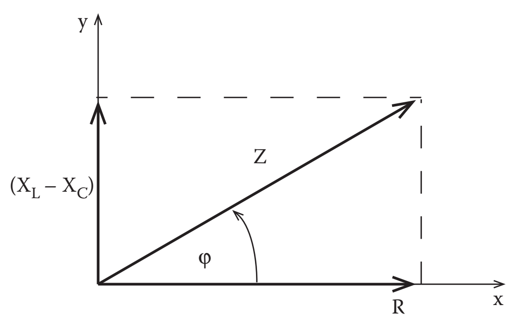


FIGURE 18.14 A diagram, geometrically similar to Figure 18.13, gives the relation for total circuit impedance.

developed for R, C, and L in series. The equations can be modified for a simple series RC or RL circuit by setting the appropriate voltage (V_{0L} or V_{0C}) and reactance (X_L or X_C) equal to zero.

EXAMPLE 18.2

Consider an RCL series circuit (Figure 18.11), with $V_{0S} = 170.0$ V, $\omega = 377$ rad/s, $R = 100.0\ \Omega$, $C = 8.00\ \mu F$, and $L = 0.400$ H. Determine numerical values for (a) X_C, (b) X_L, (c) Z, (d) I_0, (e) φ, and (f) V_{0R}, V_{0C}, and V_{0L}.

Solution

a. $X_C = \dfrac{1}{\omega C} = \dfrac{1}{(377 \text{ rad/s})(8.00 \times 10^{-6}\text{ F})} = 332\ \Omega.$

b. $X_L = \omega L = (377 \text{ rad/s})(0.400\text{ H}) = 151\ \Omega.$

c. $Z = \sqrt{R^2 + (X_L - X_C)^2} = \sqrt{(100.0\ \Omega)^2 + (151\ \Omega - 332\ \Omega)^2} = 207\ \Omega.$

d. $I_0 = \dfrac{V_{0S}}{Z} = \dfrac{170.0\text{ V}}{207\ \Omega} = 0.821$ A.

e. $\varphi = \tan^{-1}\dfrac{(X_L - X_C)}{R} = \tan^{-1}\dfrac{(151\ \Omega - 332\ \Omega)}{100.0\ \Omega} = \tan^{-1}(-1.21) = -50.4°.$

f. $V_{0R} = I_0 R = (0.821\text{ A})(100.0\ \Omega) = 82.1$ V,
$V_{0C} = I_0 X_C = (0.821\text{ A})(332\ \Omega) = 273$ V,
$V_{0L} = I_0 X_L = (0.821\text{ A})(151\ \Omega) = 124$ V.

EXAMPLE 18.3

Simple RC circuit. Suppose the inductor in Figure 18.11 is replaced with a straight conducting wire. That is, the inductor is removed from the circuit. If all other parameters have the same values as in Example 18.2, determine numerical values for (a) X_C, (b) Z, (c) I_0, (d) φ, and (e) V_{0R} and V_{0C}.

Solution

a. $X_C = \dfrac{1}{\omega C} = \dfrac{1}{(377 \text{ rad/s})(8.00 \times 10^{-6}\text{ F})} = 332\ \Omega.$

b. $Z = \sqrt{R^2 + (-X_C)^2} = \sqrt{(100.0\ \Omega)^2 + (-332\ \Omega)^2} = 347\ \Omega.$

c. $I_0 = \dfrac{V_{0S}}{Z} = \dfrac{170.0\text{ V}}{347\ \Omega} = 0.490$ A.

d. $\varphi = \tan^{-1}\dfrac{(-X_C)}{R} = \tan^{-1}\dfrac{(-332\ \Omega)}{(100.0\ \Omega)} = \tan^{-1}(-3.32) = -73.2°.$

e. $V_{0R} = I_0 R = (0.490\text{ A})(100.0\ \Omega) = 49.0$ V,
$V_{0C} = I_0 X_C = (0.490\text{ A})(332\ \Omega) = 163$ V.

EXAMPLE 18.4

Simple RL circuit. Suppose the capacitor in Figure 18.11 is replaced with a straight conducting wire and all other parameters have the values stated in Example 18.2. Determine numerical values for: (a) X_L, (b) Z, (c) I_0, (d) φ, and (e) V_{0R} and V_{0L}.

Solution

a. $X_L = \omega L = (377 \text{ rad/s})(0.400\text{ H}) = 151\ \Omega.$

b. $Z = \sqrt{R^2 + X_L^2} = \sqrt{(100.0\ \Omega)^2 + (151\ \Omega)^2} = 181\ \Omega.$

c. $I_0 = \dfrac{V_{0S}}{Z} = \dfrac{170.0\text{ V}}{181\ \Omega} = 0.939$ A.

d. $\varphi = \tan^{-1}\left(\frac{V_L}{Z}\right) = \tan^{-1}\left(\frac{151\ \Omega}{100.0\ \Omega}\right) = \tan^{-1}(1.51) = 56.5°.$

e. $V_{0R} = I_0R = (0.939\ \text{A})(100.0\ \Omega) = 93.9\ \text{V},$
$V_{0L} = I_0X_L = (0.939\ \text{A})(151\ \Omega) = 142\ \text{V}.$

18.4.1 RCL Series Resonance

In ac circuits, the quantities X_C and X_L and correspondingly I_0, V_{0C}, and V_{0L} depend on the value of the source frequency $\omega = 2\pi f$. It is reasonable to ask "Is there a frequency at which, for a fixed value of V_{0S}, I_0 is a maximum?" The answer is "yes" and that frequency (f_R) is called the *resonant frequency*. The concept of a quantity being enhanced, or going to a maximum value, is called *resonance*. In this case, the quantity is I_0. To determine f_R, write Equation 18.20 as

$$I_0 = \frac{V_{0S}}{Z} = \frac{V_{0S}}{\sqrt{R^2 + (X_L - X_C)^2}}. \tag{18.23}$$

For a fixed value of V_{0S}, I_0 will have a maximum value when Z is a minimum. Since R does not depend on ω (except for extremely high values of ω), Z will be a minimum when $(X_L - X_C) = 0$ or $X_L = X_C$. Thus,

$$\omega_R L = \frac{1}{\omega_R C}$$

or

$$\omega_R^2 = \frac{1}{LC} = (2\pi f_R)^2.$$

So,

$$f_R = \frac{\omega_R}{2\pi} = \frac{1}{2\pi\sqrt{LC}}. \tag{18.24}$$

EXAMPLE 18.5

Determine f_R for the series circuit in Example 18.2. Here, of course,

$$\omega \neq 377\ \text{rad/s}.$$

Solution

Equation 18.24 gives

$$f_R = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{(0.400\ \text{H})(8.00\times 10^{-6}\ \text{F})}} = 89.0\ \text{Hz}.$$

18.5 ROOT-MEAN-SQUARE AVERAGE VALUES, POWER

In the previous sections, time-varying sinusoidal voltages and currents were discussed. Since these vary from a maximum amplitude through zero to a negative amplitude, it is reasonable to define an average value for them. To determine the average value of something, as an example, the average age of people in a classroom, their ages would be added up and the sum divided by the number of people in the room. To determine the average value of a function, in an interval of its variable, the function would be "added up" (i.e., integrated over its variable via calculus) and divided by the interval. If this is done for a sinusoidal function, over one period, the average is zero. This result is reasonable since the function is negative as much as it is positive, so adding its contributions to the sum (integral) should give zero. Consider a different averaging process. To circumvent the negative contributions, the sinusoidal function is first squared, then averaged over one period. This would give the average of the squared function. For $V = V_0 \sin \omega t$, the result is

$$(V^2)_{ave} = \frac{V_0^2}{2}.$$

The final step, taking the square root of this average, yields the root of the mean of the square, that is, the *root-mean-square* (rms) value. Thus,

$$V_{rms} = \sqrt{(V^2)_{ave}} = \sqrt{\frac{V_0^2}{2}} = \frac{V_0}{\sqrt{2}} = 0.707\ V_0. \tag{18.25}$$

Similarly, for a sinusoidal current

$$I_{rms} = \frac{I_0}{\sqrt{2}} = 0.707\ I_0. \tag{18.26}$$

The rms values are called "effective" values. In particular, a sinusoidal current with a given I_{rms} creates the same amount of heating in a resistor as does a dc current of the same value, that is,

$$P = (I_{rms})^2 R = (I_{dc})^2 R. \tag{18.27}$$

The power delivered to a circuit or a single circuit element is equal to the product of the current through and voltage across the circuit/element. But, unlike the dc case, I and V are not in phase, that is, in general $\varphi \neq 0$, so the instantaneous power is

$$P = IV = (I_0 \sin \omega t)[V_0 \sin(\omega t + \varphi)].$$

When Equation 18.27 is averaged over one cycle, the result is

$$P = \left(\frac{I_0 V_0}{2}\right)\cos\varphi = \left(\frac{I_0}{\sqrt{2}}\right)\left(\frac{V_0}{\sqrt{2}}\right)\cos\varphi = I_{rms}V_{rms}\cos\varphi. \tag{18.28}$$

Note: It is conventional to omit the subscript "rms" on I, V, and P. For ac circuits, it is understood that the specified quantities are rms values.

EXAMPLE 18.6

Refer to Example 18.2. Determine (a) the numerical value of the power delivered to the RCL series circuit by the source and (b) the rms voltage across the resistor.

Solution

a. Using Equation 18.28

$$P = \left(\frac{I_0 V_{0S}}{2}\right)\cos\varphi = \frac{(0.821\ \text{A})(170.0\ \text{V})}{2}\cos(-50.4°) = 44.5\ \text{W}.$$

b. $V_{OR} = (I_0/1.41)R = (0.821\text{A}/1.41)\ (100\ \Omega) = 58.2\ \text{V}$

PROBLEMS

18.1 An RCL series circuit is driven by a source at a frequency of 60.0 Hz. The circuit resistance is 150.0 Ω, $X_C = 280.0\ \Omega$, and the inductance is L = 0.293 H. Determine the value of the circuit impedance Z.

18.2 A resistor R = 300.0 Ω, a capacitor $C = 1.20 \times 10^{-6}$ F, and an inductor L = 14.0 H are connected in series to a sinusoidal voltage source with amplitude $V_{0S} = 120.0$ V. Determine the resonant linear frequency of this circuit.

18.3 An RCL series circuit is driven by a source voltage given by $V = V_{0S}\sin(\omega t + \phi)$. Here $V_{0S} = 150.0$ V, $\omega = 120\pi$ rad/s, R = 750.0 Ω, $X_L = 320.0\pi\ \Omega$, and $X_C = 1250\ \Omega$. Determine numerical values for (a) the circuit impedance, (b) the current amplitude, (c) V_{0C}, (d) the phase angle ϕ, (e) the inductance L, (f) the capacitance C, (g) the resonant frequency f, and (h) the average power supplied to the circuit by the generator.

18.4 A simple armature consists of three similar rectangular loops of wire that are 0.180 m long and 0.080 m wide. The coplanar loops are immersed in a spatially uniform, 0.170 T magnetic field. (a) At what constant angular speed must this armature rotate in order to have an induced emf of 13.0 V? (b) Determine the frequency of this induced voltage.

18.5 An RCL series circuit is driven by a source with voltage amplitude 120.0 V. The circuit elements have values R = 90.0 Ω, $X_C = 20.0\ \Omega$, and $X_L = 50.0\ \Omega$. Determine (a) The circuit current and (b) the power dissipated by the resistor, capacitor, and inductor.

18.6 An RCL series circuit is driven by a sinusoidal voltage source of constant amplitude 170.0 V. If ω = 377 rad/s, L = 4.00 H, C = 3.00 μF, and R = 1200.0 Ω, calculate (a) X_L, (b) X_C, (c) Z, (d) I_0, (e) I_{rms}, (f) $(V_L)_{rms}$, and (g) f_R.

18.7 A capacitor in an ac circuit has a capacitive reactance of 26.0 Ω at a frequency of 120.0 Hz. Determine X_C for this capacitor when the circuit frequency is 340.0 Hz.

18.8 An 85.0-Ω resistor is connected, in series with a capacitor, in an ac circuit. The circuit voltage is $V = (120.0)\sin\omega t$. At a given source frequency, the capacitive reactance is found to be $X_C = 50.0\ \Omega$. Determine (a) the phase angle φ and (b) the circuit current amplitude.

18.9 The inductive reactance of an inductor is $X_L = 370.0\ \Omega$ when the frequency is 60.0 Hz. Determine its reactance at a frequency of 280.0 Hz.

18.10 Determine the frequency (in Hz) at which the reactances of an L = 42.0 mH inductor and a C = 74.0 μF capacitor are equal.

18.11 A 100.0-Ω resistor is connected, in series with an inductor, in an ac circuit. The circuit voltage source is $V = (170.0)\sin\omega t$. At a given source frequency, the inductive reactance is found to be $X_L = 26.0\ \Omega$. Determine (a) the phase angle φ and (b) the circuit current amplitude.

18.12 The power dissipated in an RCL series circuit is 75.0 W. The current and voltage amplitudes are $I_0 = 2.65$ A and $V_0 = 170.0$ V, respectively. Calculate the power factor angle φ for this circuit.

18.13 The resonant frequency of a series RCL circuit is $f_r = 2.60$ kHz. Determine the new resonant frequency if the inductance is tripled and the capacitance is doubled.

18.14 An RCL series circuit is powered by a source with a voltage amplitude $V_{0S} = 170.0$ V. The circuit elements have values of $R = 100.0\ \Omega$, $X_C = 30.0\ \Omega$, and $X_L = 55.0\ \Omega$. Determine (a) the circuit current, and (b) the power dissipated separately by the resistor, capacitor, and inductor.

18.15 An ac generator with output $V = (18.0)\sin \omega t$ at frequency $f = 500.0$ Hz is connected in series with a 22.0-Ω resistor, a 6.20-μF capacitor, and a 7.36 mH inductor. Determine the rms voltage across each element.

19 Electromagnetic Waves

19.1 INTRODUCTION

In 1865, the great Scottish physicist, James Clerk Maxwell (1831–1879), predicted the existence of electromagnetic waves (EMWs). His prediction led to the development of wireless communication and to a deeper understanding of a large range of physical phenomena. It is now understood that EMWs are generated when (a) charged particles are accelerated, (b) electrons in atoms make transitions from one atomic energy state to another, (c) nuclear particles make transitions from one nuclear energy state to another, or (d) nuclear particles disintegrate from one form to another. These ideas were not known in Maxwell's time, but he made his ingenious discovery by investigating the known equations of electricity and magnetism. His exact derivation is beyond the level of this text, but consider the following simplified analysis.

Faraday's law is

$$\xi_i = -\frac{\Delta\Phi_B}{\Delta t} = \sum E \cdot \Delta l. \tag{19.1}$$

A charged particle in the region where $(\Delta\Phi_B/\Delta t) \neq 0$ would experience an induced electromotive force (emf) ξ_i and thus an electric field, as the right side of Equation 19.1 indicates. So a time-varying **B** field gives rise to an **E** field. Maxwell observed that Ampere's law was incomplete insofar as it did not allow for the charge continuity equation to hold. He proposed that Ampere's law should include a term added to it as shown in Equation 19.2.

$$\sum_{\text{A.L.}} B \cdot \Delta l = \mu_0 \left(I + \varepsilon_0 \frac{\Delta\Phi_E}{\Delta t} \right). \tag{19.2}$$

Equation 19.2 indicates that a **B** field can be created by a current (I) or by the added term $\mu_0\varepsilon_0(\Delta\Phi_E/\Delta t)$. The added term contains a time-varying electric field, so Equation 19.2 implies that a changing **E** field gives rise to a **B** field. Maxwell manipulated Equations 19.1 and 19.2 and with the use of some vector identities, acquired separate differential equations for **E** and **B**. These equations were recognized as wave equations whose solutions are waves (oscillations) of **E** and **B** propagating through space and varying both spatially and temporally. So, time-varying **E** fields create time-varying **B** fields and vice versa.

To understand the nature and properties of EMWs, it is instructive to analyze the production of radio waves from a simple oscillating source connected to a dipole antenna (Figure 19.1). The two rods are collinear conductors with a very small gap between them. The ac generator produces a sinusoidal emf, current, and charge distribution between the terminals A and C that are connected to the metal rods that constitute the antenna (Figure 19.2).

The antenna forms an open circuit. The generator and antenna form a simple EMW transmitter. Recall that separation of charges gives rise to **E** fields and currents create **B** fields. Figure 19.3a through d depict "snapshots" of the field patterns that correspond to charge regions in Figure 19.2.

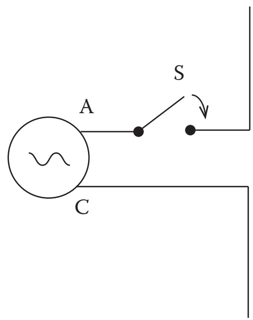


FIGURE 19.1 A simple dipole antenna.

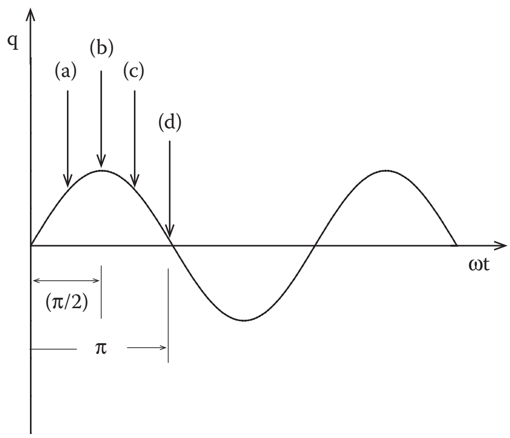


FIGURE 19.2 The charge distribution on the antenna as a function of time.

19.1.1 Analysis

1. In region (a), the current is increasing and moving upward, so the upper rod has an excess of positive charge, the lower rod a deficiency. The **E** field is directed from positive charges to negative and is perpendicular to the antenna at the location of the charges. The current is flowing upward, so the **B** field, from right hand rule-II, forms counterclockwise concentric circles, as viewed from above about the vertical axis of the antenna.
   Note: For clarity, the generator is not shown, and only those field components in the plane of the figure and to the right of the antenna will be displayed. In reality, these fields totally surround the antenna.
2. At point (b), which corresponds to $\omega t = \pi/2$, the charge on the upper rod is a maximum and the current $I = \Delta q/\Delta t$ is zero. So, $\mathbf{B} = 0$ and the **E** field generated this instant, and depicted as closest to the antenna, has no **B** field shown. The **E** and **B** fields, produced earlier in the cycle, that is, when $\omega t < \pi/2$, however, have not disappeared. Recall that the fields separately have energy densities $(1/2)\varepsilon_0 E^2$ and $(1/2\ \mu_0)B^2$. This field energy is propagated away from the antenna at the speed of light propagation velocity.
3. In region (c), the current is now flowing downward and the fields have reversed their directions. The directions of **E** and **B** that were created earlier in the cycle are not affected by later changes in antenna current and charge distribution.
4. At point (d), which corresponds to $\omega t = \pi$, there is no net charge between the antenna rods, so the **E** field near the antenna is zero. The fields created earlier in the cycle ($\omega t < \pi$) continue to propagate away from the antenna. The propagated fields vary both spatially and

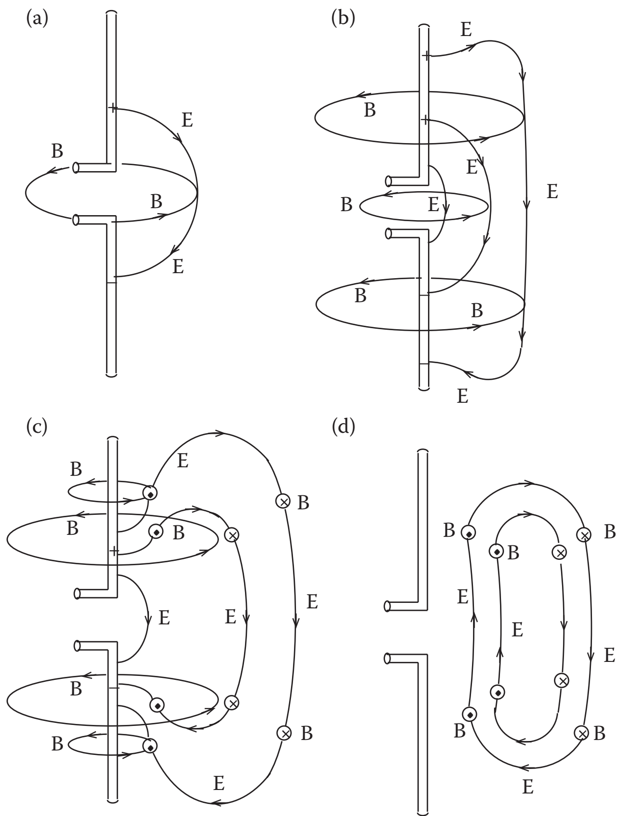


FIGURE 19.3 The electric and magnetic field lines near an antenna and the creation of EMWs.

temporally (with time). At a fixed spatial point, the fields vary temporally. At a fixed temporal point (snapshot), the spatial variation of the fields may be observed. The field patterns in Figure 19.3d may not be intuitive. We know that lines of **B** form closed loops around the antenna. The lines of **E** closer to the antenna are directed upward, while the more distant lines point downward. Traversing from left to right in the "snapshot" of Figure 19.3d, the lines of **E** can reverse direction if they form closed loops as shown. During the second half of the charge cycle (Figure 19.2), the fields of Figure 19.3d continue to propagate outward, while a new set of lines, with field directions reversed, are formed at this later time.

A rigorous treatment of the dipole antenna, as opposed to the simplified treatment just given, shows that the "near field," that is, the field near the antenna, is that of an oscillating electric dipole, but the "far field" is similar to Maxwell's solutions of the **E** and **B** wave equations. In this and the following two chapters, we will be interested in the "far-field" EMWs.

19.1.2 Properties of EMWs

1. **E** and **B** are mutually perpendicular, as shown in Figure 19.3.
2. Maxwell showed that **E** and **B** are in phase, that is, **E** and **B** reach their maximum values at the same time.
3. Maxwell also predicted that the fields propagate at a velocity $v = (1/\sqrt{\varepsilon_0\mu_0}) = c$. That is, EMWs propagate at $c \approx 3 \times 10^8$ m/s, which equals the speed of light in a vacuum or air. The properties of EMWs are depicted in Figure 19.4.

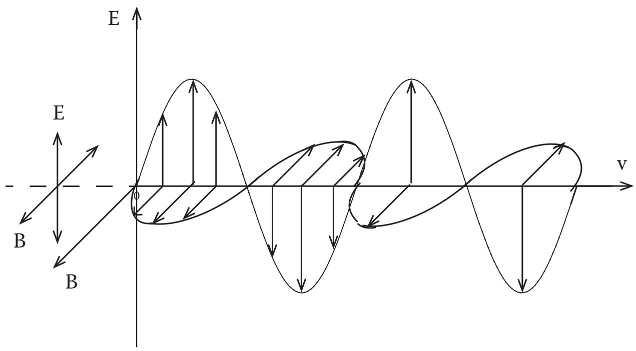


FIGURE 19.4 In a plane EMW, the **E** and **B** oscillations are in phase and perpendicular to each other.

Some of the parameters used to characterize waves in general are wavelength (λ), frequency (f), velocity (v), and amplitude (E_0) (Figure 19.5).

A wave is a periodic disturbance in a medium. In the case of EMWs, the disturbance in the oscillations and the media are **E** and **B** fields. The parameters are defined as

a. Wavelength (λ) = spatial extent of one complete oscillation. The units are meters.
b. Frequency (f) = the number of complete oscillations per second. Units are oscillations/second = Hertz (Hz).
c. Velocity (**v**) = the rate of propagation of the disturbance (oscillation). Units are meter per second.
d. Amplitude ($\mathbf{E_0}$ or $\mathbf{B_0}$) = the maximum extent of the oscillations. Their units are equal to the units of the oscillation.

For EMWs, E_0 = N/C, B_0 = Tesla.

The solution to Maxwell's wave equations, in one dimension, can be written as

$$E = E_0 \sin(kx - \omega t), \tag{19.3}$$

$$B = B_0 \sin(kx - \omega t), \tag{19.4}$$

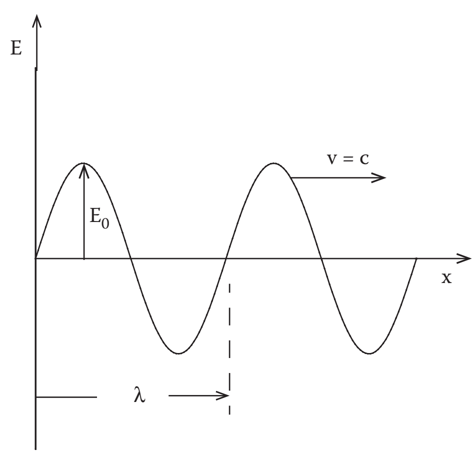


FIGURE 19.5 In an EMW, the **E** and **B** fields are the medium and their oscillations are the wave.

where the relations $k = 2\pi/\lambda$ and $\omega = 2\pi f$ have been used. These are traveling wave equations and describe both the spatial and temporal variations of the **E** and **B** fields. EMWs carry energy. The energy per unit area per second (i.e., the energy intensity) crossing a plane perpendicular to the velocity and averaged over one cycle can be shown to be

$$S_{ave} = \frac{1}{2\mu_0}(E_0 B_0). \tag{19.5}$$

The units of S_{ave} are watts/m². Also, the electric field is related to the B field by the relation

$$E = cB. \tag{19.6}$$

An operational definition of the magnitude of the velocity of EMWs can be obtained from the mechanical definition

$$v = \frac{\text{displacement}}{\text{time elapsed}}.$$

From Figure 19.3, the time required for the disturbance to travel a distance of one wavelength λ is equal to the period of that oscillation. Thus,

$$v = \frac{\lambda}{\Gamma} = \lambda\left(\frac{1}{\Gamma}\right) = \lambda f. \tag{19.7}$$

EXAMPLE 19.1

AM radio station KDKA (Pittsburgh) broadcasts at a power of 50,000 W. Determine the electric field amplitude E_0 of its signal at a distance $r = 10.0$ miles from its antenna.

Solution

$$S_{ave} = \frac{P_{ave}}{A} = \frac{1}{2\mu_0}(E_0 B_0) = \frac{1}{2\mu_0} E_0\left(\frac{E_0}{c}\right). \tag{19.8}$$

If the power is transmitted isotropically, the surface area of a sphere of radius $r = 10.0$ miles = 16.1 km is relevant. Rearranging Equation 19.8

$$E_0^2 = \frac{2\mu_0 c P_{ave}}{4\pi r^2} = \frac{2(4\pi \times 10^{-7}\ \text{Tm/A})(3.00 \times 10^8\ \text{m/s})(5.00 \times 10^4\ \text{W/m}^2)}{4\pi(1.61 \times 10^4\ \text{m})^2},$$

$$E_0^2 = 1.14 \times 10^{-2}\ \text{V}^2/\text{m}^2,$$

and

$$E_0 = 0.107\ \text{V/m}.$$

This is a fairly large radio signal.

19.2 ELECTROMAGNETIC SPECTRUM

The traveling waves of Equations 19.3 and 19.4 are solutions to Maxwell's wave equations regardless of their frequency or wavelength. Thus, all EMWs have the same basic character regardless

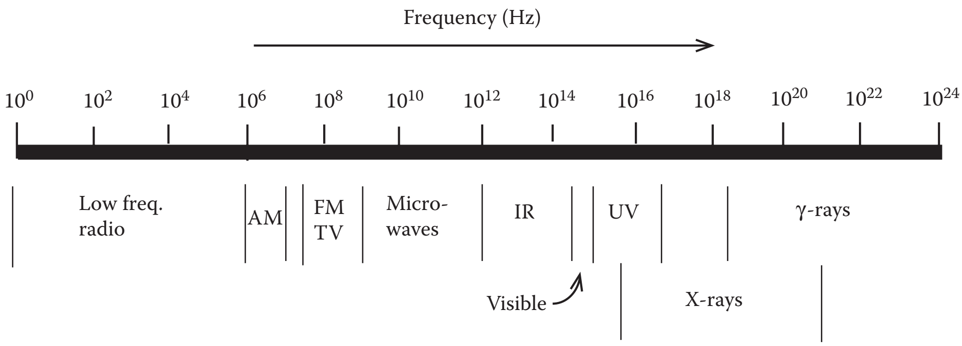


FIGURE 19.6 The approximate frequency ranges for the EMW spectrum.

of their frequency. It is conventional to assign different names to the waves in various frequency regions (bands) (see Figure 19.6).

For most of the frequency bands in Figure 19.6, the endpoints are not well defined and there is some overlap, with the exception of the AM and FM-TV bands that are defined legally. Notice the narrowness of the visible band which extends from approximately 4.29×10^{14} to 7.50×10^{14} Hz. Within this band, the colors of visible light depend on the wavelength of the EMW.

EXAMPLE 19.2

Determine (a) the wavelength of the corresponding endpoints of the visible spectrum with frequencies of $f_1 = 4.29 \times 10^{14}$ Hz and $f_2 = 7.50 \times 10^{14}$ Hz, and (b) the wavelength of the carrier signal of AM radio station KDKA that broadcasts at f = 1020 kHz.

Solution

a. Since $v = c = \lambda f$,

$$\lambda_1 = \frac{c}{f_1} = \left(\frac{3.00 \times 10^8 \text{ m/s}}{4.29 \times 10^{14} \text{ Hz}}\right) = 6.99 \times 10^{-7} \text{ m} = 699 \text{ nm}.$$

Also

$$\lambda_2 = \left(\frac{3.00 \times 10^8 \text{ m/s}}{7.50 \times 10^{14} \text{ Hz}}\right) = 4.00 \times 10^{-7} \text{ m} = 400 \text{ nm}.$$

b. $\lambda_{KDKA} = \dfrac{c}{f} = \dfrac{3.00 \times 10^8 \text{ m/s}}{1.02 \times 10^6 \text{ Hz}} = 294 \text{ m}.$

So, one wavelength of KDKA's signal is approximately the length of three football fields.

PROBLEMS

19.1 EMWs result whenever
- a. A separation of charge is produced
- b. Unbalanced charges are in motion
- c. Electric or magnetic fields are established in space
- d. Unbalanced charges are accelerated and/or particle energy transitions occur in atoms
- e. All of the above

19.2 The frequency range of UHF television is from 470.0 MHz (channel 14) to 890.0 MHz (channel 83). Determine the wavelength range.

19.3 A certain EMW has a frequency of 1.02×10^6 Hz, wavelength equal to 294 m, and an electric field amplitude $E_0 = 1.90 \times 10^{-2}$ V/m. Calculate numerical values for its wave number (k), angular frequency (ω), and magnetic field amplitude B_0.

19.4 Radio station KDKA-AM in Pittsburgh, Pennsylvania, is advertised as a 50,000-W station. Assume that it transmits that power isotropically over a hemisphere. Determine the amplitude of the electric field detected from KDKA by a radio receiver located at 100.0 km from Pittsburgh.

19.5 A radio station transmits an EMW that has an amplitude $E_0 = 1.73 \times 10^{-2}$ V/m at a distance r = 5.00 km from the station. Calculate the average power per area received at this 5.00 km location. Again, assume that the power is transmitted isotropically over a hemisphere.

19.6 The average distance from the earth to the moon is 238,000 miles and from earth to the sun is 93,000,000 miles. Calculate the time for light to travel (a) from the moon to the earth, (b) from the sun to the earth.

19.7 The frequency range of x-rays is approximately 7.50×10^{16} Hz to 7.00×10^{19} Hz. Calculate the wavelength range.

19.8 Radio station KDKA-AM broadcasts at a carrier frequency of 1020 kHz. The value of the inductor, in its tuning circuit, is 9.67×10^{-4} H. What must be the capacitance value, in the tuning circuit, to "tune" the radio to this station?

19.9 Alpha Centauri, the second closest visible star to earth (our sun is the closest star), is approximately 4.3 light-years away. Convert that distance to meters. (*Hint*: One light-year is the distance traveled by light, in a vacuum, in 1 year.)

19.10 The average energy intensity of sunlight at the top of the earth's atmosphere is approximately 1.390 kW/m^2. If all this energy passed through our atmosphere unattenuated, (a) how much power would be delivered to a rectangular roof panel measuring 15.0 m by 6.00 m (typical area of the south half of a home roof)? (b) How much energy would be delivered in 1 h?

19.11 The average energy intensity of sunlight at the top of the earth's atmosphere is approximately 1.390 kW/m^2. For these EMWs, calculate the magnitude of (a) the amplitude of the electric field (E_0), (b) the amplitude of the magnetic field (B_0), and (c) the maximum electric force that a free electron, in the field, would experience.

19.12 The average energy intensity of solar radiation reaching the earth's surface is 920 W/m^2 (its 1390 W/m^2 at the top of the atmosphere). (a) Determine the area of a solar panel (assuming 100% conversion) needed to furnish the daily electrical energy needs (55.0 kW h) of a typical household. (b) At the very best, under ideal conditions, solar cell conversion of light to electrical energy approaches 20.0%. Recalculate the area needed in part (a), with 20.0% of 920 W/m^2.

19.13 Envision a circular loop of radius r = 8.00 cm whose plane is perpendicular to a spatially uniform electric field. The electric field is increasing at the rate $(\Delta E/\Delta t) = 100.0$ V/s m. Use Ampere's law to calculate the magnitude of the induced magnetic field on the circumference of the loop.

19.14 The electric field component of a traveling EMW is described by $E = E_0 \sin(kx - \omega t)$. Radio station KDKA-AM broadcasts at a frequency of 1020 kHz. Calculate the instantaneous value of E at a position x = 100.0 m from its broadcast antenna, where $E_0 = 1.95$ mV/m and t = 30.0 s from initiation of the signal.

19.15 At a given spatial point, what values of the electric and magnetic field amplitudes, E_0 and B_0, are required to produce an average energy intensity of 920.0 W/m^2?

20 Geometrical Optics

20.1 INTRODUCTION

The electromagnetic spectrum spans more than 24 orders of magnitude with frequencies from <10 to >10^{24} Hz. The human body has senses that detect only a tiny portion of that spectrum. The molecules in our skin detect a small portion of the infrared (IR) range. Our skin feels warm or may "burn" if we are outside in bright sunlight. Healthy human eyes can detect frequencies in the range of 4.29×10^{14} to 7.50×10^{14} Hz. This is the *visible range* and the corresponding electromagnetic waves (EMWs) are called visible light. This chapter and the next will address the properties and some phenomena associated with visible light. *Geometrical optics* addresses the interaction of light with obstacles whose dimensions are much larger than the wavelength of light.

20.2 WAVEFRONTS AND RAYS

Two useful concepts in the discussion of light are wavefronts and rays. An isolated point source of light gives off EMWs in all directions. For clarity in Figure 20.1, only three waves, emitted to the right, are shown, but one can envision a large number emitted simultaneously. Recall that EMWs vary both spatially and temporally. Figure 20.1 depicts a "snapshot" of the emitted light waves. If points of like phase, for example, the first maxima of the waves are joined by an imaginary line, the locus of connected points forms an imaginary sphere (or a circle in two dimensions). In Figure 20.1, the smallest semicircle of radius r_3 represents the connection of the maxima closest to the source at the time the snapshot was taken.

These connected points are called *wavefronts* and move outward from the point source. The wavefronts in Figure 20.1 are *spherical wavefronts* and the associated waves are called *spherical waves*. In an isotropic medium, the velocity of the EMW is perpendicular to the wavefronts.

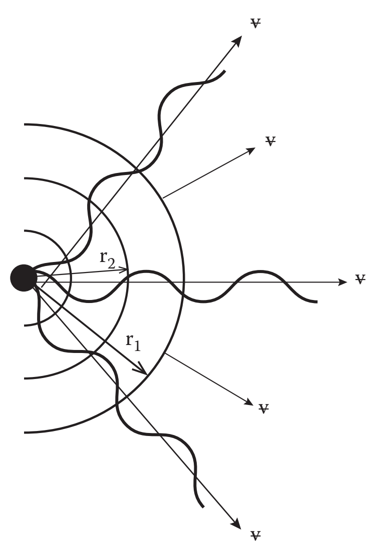


FIGURE 20.1 Wavefronts are the locus of points of like phase.

Vectors drawn perpendicular to the wavefronts and parallel to the velocity vectors are called *rays* (see Figure 20.2). Note that at great distances from the point light source, any small segment of the spherical wavefront approximates a plane wavefront. The associated waves are called *plane waves* and their rays are parallel to each other.

One additional concept, Fermat's principle, will prepare us for discussion of the two topics of this chapter, namely reflection and refraction. *Fermat's principle of least time* (Pierre de Fermat, French mathematician, 1601–1665) states:

> In traveling from one point to another, a light ray will traverse the path for which the propagation time has a minimum value.

Notice that it is the time and not necessarily the distance that is a minimum. Suppose an equation for the time required for light to travel from a given point, interact with an obstacle (e.g., mirror surface and transparent solid) and arrive at a final point, is constructed. If such an equation is minimized (via calculus) with respect to time (Fermat's principle), two important optical laws are derived. They are *the law of reflection* and *the law of refraction* (Snell's law).

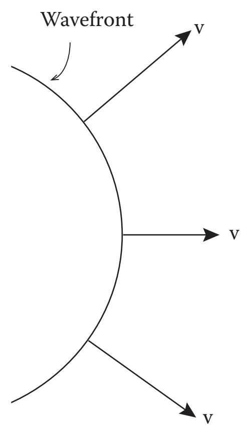


FIGURE 20.2 In an isotropic medium, rays are perpendicular to wavefronts.

20.3 REFLECTION

Incident rays of light that encounter an obstacle such as a sheet of paper or a piece of shiny aluminum and are deflected are said to be "reflected."

20.3.1 Law of Reflection

A beam or a ray of light from a laser is shined on a mirror at an incident angle θ_i, measured with respect to a line drawn perpendicular (normal) to the mirror surface (Figure 20.3). The reflected beam makes an angle θ_r with the normal. Fermat's principle gives: the *angle of incidence* equals *The angle of reflection*. That is,

$$\theta_i = \theta_r \text{ (law of reflection).} \tag{20.1}$$

Reflections are categorized as either *specular* or *diffuse*. Specular is reflection from an object whose surface irregularities are small compared with the wavelength of light. That is, very "smooth" surfaces, such as mirrors and polished metals (Figure 20.3).

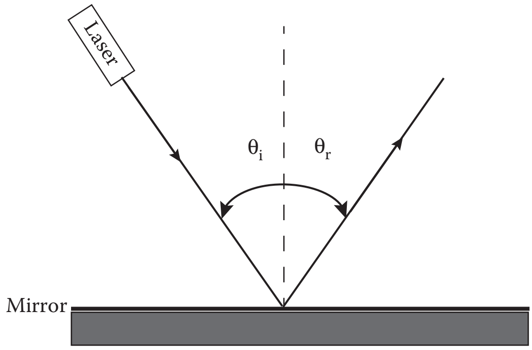


FIGURE 20.3 For specular reflection, the angles of incidence and reflection are equal.

Diffuse reflection occurs when surface irregularities or roughness are equal or greater than the wavelength of light (Figure 20.4). Examples of such surfaces are most paper, manila folders unpolished metals, and objects painted with "flat" paint. For rough surfaces, the law of reflection applies to each point on the surface. If a large number of parallel rays are incident on such a surface, the reflected rays will not be parallel, but will propagate from the surface in many different directions. For the remainder of this chapter, only specular reflections will be considered.

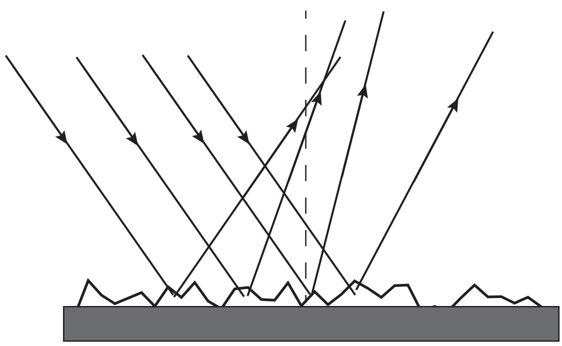


FIGURE 20.4 An illustration of diffuse reflection.

20.3.2 Plane Mirrors

Figure 20.5 depicts a small light source at the top of a candlestick located a distance s_0 from a plane mirror. In optics jargon, a source of light is called an *object*, so s_0 is the *object distance*. The source emits light rays in all directions and some are reflected by the mirror. Consider two such rays labeled "1" and "2" in Figure 20.5. For each ray $\theta_i = \theta_r$. After reflection, the rays diverge as shown. If you place your eye, as shown in Figure 20.5, the diverging rays will enter it and the eye will see an *image* of the light source. If the extensions of the reflected rays are drawn behind the mirror, they will intersect in a plane. This plane is called the *image plane* and its distance from the mirror is called the *image distance*, which is labeled s_i.

Physiologically, the eye interprets all rays that enter it as traveling along straight lines, so the image will appear to be behind the mirror. In Figure 20.5, triangle (abd) is similar to triangle (cbd). Since side (bd) is common to both, the triangles are also equal. Thus, for a plane mirror

$$-s_0 = s_i. \tag{20.2}$$

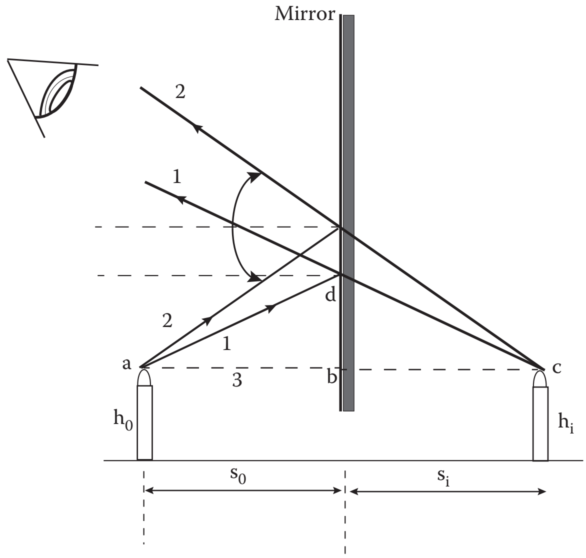


FIGURE 20.5 Reflection from a plane mirror.

That is, the image distance is equal to the negative of the object distance. Note that these distances are measured from the mirrored surface. A sign convention will be discussed in a later section. Additionally, ray "3" and its extension are along a line perpendicular to the mirror surface, so here,

$$h_i = h_0. \tag{20.3}$$

In order to compare image size to object size, it is conventional to define a quantity called the *lateral magnification* M with magnitude

$$|M| \equiv \frac{h_i}{h_0}. \tag{20.4}$$

Note: Although we have addressed the image of a small light source (Figure 20.5), the object need not emit light. We "see" objects either because they emit light or when light is reflected from them. So the object, in the above treatment, need not be a light-emitting source itself.

20.3.3 Spherical Mirrors

If a large hollow sphere of radius (r) is polished or "silvered" on its *inner* surface so that it is a reflector, and a small circular segment is cut out of its surface, that segment is a *concave spherical mirror*. Alternatively, if the sphere is polished on its *outer* surface and a segment is removed, that segment is a *convex spherical mirror* (see Figure 20.6).

Like plane mirrors, spherical mirrors can be used to form images of a light-reflecting or light-emitting source. To analyze these images, the concepts of focal point and focal length are needed. Consider the parallel rays, from a distant source, that are reflected from a spherical mirror (Figure 20.7). Wherever on the mirror surface the rays are reflected, the law of reflection applies. Also, the normals to the surface are the radii emanating from the center of curvature (C).

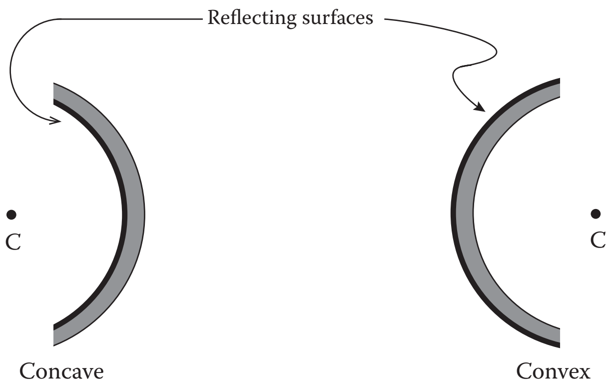


FIGURE 20.6 Reflection from concave and convex spherical surfaces.

For the concave mirror (Figure 20.7a), the reflected rays all converge and pass through a common point (F) on an axis called the *principal axis* passing through the vertex (V). (F) is called the *focal point* and its distance from (V) is called the *focal length* and designated by (f). Rays of light, carrying EMW energy, pass through the focal point.

For the convex mirror (Figure 20.7b), the reflected rays all diverge and do not pass through a common point. However, to an observing eye, the reflected rays would *appear* to emanate from a common point, the focal point (F) located *behind* the mirrored surface. For both concave and

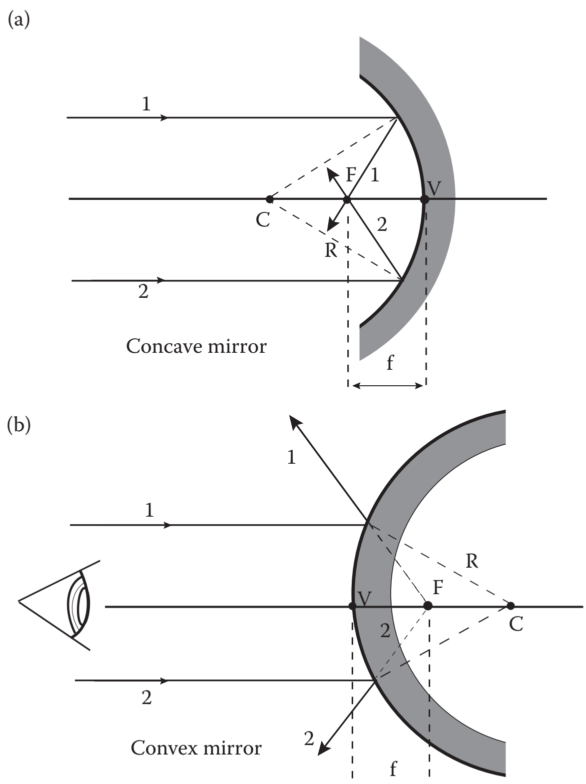


FIGURE 20.7 (a) All incident rays parallel to the principal axis pass through the focal point after reflection. (b) Parallel incident rays, after reflection, appear to emanate from the focal point.

convex mirrors, since $\theta_i = \theta_r$, the location of F is just one-half the length of the mirror radius of curvature. So,

$$|f| = \frac{R}{2}. \tag{20.5}$$

The convention is that f of a concave mirror is positive while f of a convex mirror is negative.

20.3.4 Reflected Images

Concave mirror: Consider a light-emitting or light-reflecting object of height h_0 at a distance s_0 from the vertex V of concave mirror. The size and location of the image may be determined by tracing several rays (to scale) or by an equation, here to be derived (Figure 20.8).

"Ray 1" is chosen parallel to the principal axis and, after reflection, passes through the focal point F. The reflected ray makes an angle β with the principal axis where "Ray 2" is chosen to reflect at the vertex and makes an angle α with the principal axis. After reflection, rays "1" and "2" intersect in a plane, called the *image plane*, at a distance s_i from V. Indeed, if many other rays were chosen and the law of reflection was applied to each, they would all, after reflection, intersect in the image plane. From Figure 20.8,

$$\tan\alpha = \frac{h_0}{s_0} = \frac{h_i}{s_i}. \tag{20.6}$$

Also,

$$\tan\beta = \frac{h_i}{s_i - f} \cong \frac{h_0}{f}. \tag{20.7}$$

The approximation in Equation 20.7 comes from assuming $\Delta \ll f$ so that $(f - \Delta) \cong f$. Solving both Equations 20.6 and 20.7 for (h_i/h_0) gives

$$\frac{h_i}{h_0} = \frac{s_i}{s_0} = \frac{(s_i - f)}{f} = \frac{s_i}{f} - 1. \tag{20.8}$$

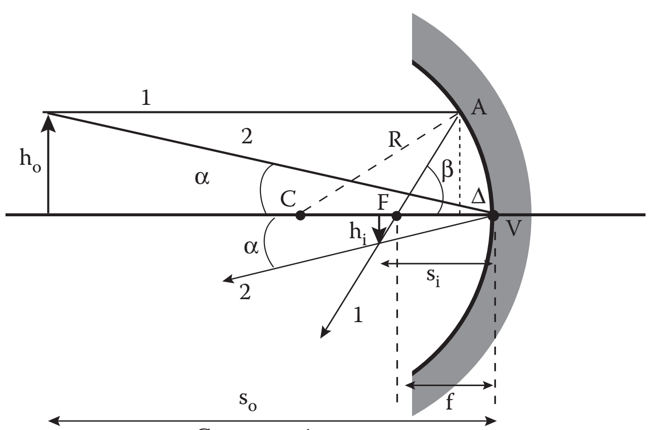


FIGURE 20.8 Rays used to locate the image formed by a concave mirror.

Dividing the second and fourth terms of Equation 20.8 by s_i and rearranging yields

$$\frac{1}{s_0} + \frac{1}{s_i} = \frac{1}{f}. \qquad \text{(mirror equation)} \qquad (20.9)$$

The utility of Equation 20.9 is that one need not construct a scale diagram to locate the image formed by a convex mirror. The image shown in Figure 20.8 is inverted. If a small screen was placed in the image plane and below the principal axis, so as not to block the light going from object to mirror, an inverted image would be seen on the screen. Light energy carried by the EMWs passes through the image plane. Such an image is called a *real image*. Correspondingly, an image formed in a plane through which light energy does not pass is called a *virtual image*. Figure 20.9 depicts a virtual image.

Notice that the object is located inside the focal length. The diverging rays entering the eye to the left of the mirror would give the impression that the light rays originated behind the reflecting surface. No light energy passes through this image plane and no image would show on a screen placed there.

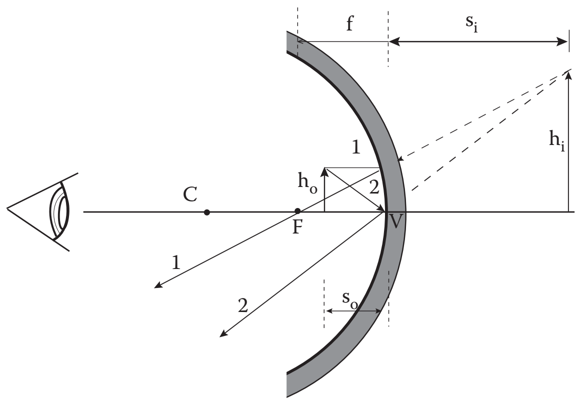


FIGURE 20.9 If the object is inside the focal length of a concave mirror, the image will be virtual.

EXAMPLE 20.1

An object of height $h_0 = 6.0$ cm is placed on the axis of a concave mirror which has a focal length of $f = 12.0$ cm. Determine the image position and magnitude of the magnification for (a) $s_0 = 18.0$ cm and (b) $s_0 = 8.0$ cm.

Solution

a. Equation 20.9 gives,

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_0} = \frac{1}{12.0\text{ cm}} - \frac{1}{18.0\text{ cm}} = \frac{3.0 - 2.0}{36.0\text{ cm}} = \frac{1.0}{36.0\text{ cm}}.$$

So,

$$s_i = 36.0\text{ cm}.$$

By Equation 20.6

$$|M| \equiv \frac{h_i}{h_0} = \frac{s_i}{s_0} = \frac{36.0 \text{ cm}}{18.0 \text{ cm}} = 2.0.$$

Thus, the image height is twice as large as the object height.

b.

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_0} = \frac{1}{12.0 \text{ cm}} - \frac{1}{8.0 \text{ cm}} = \frac{2.0 - 3.0}{24.0 \text{ cm}} = \frac{-1.0}{24.0 \text{ cm}},$$

or

$$s_i = -24.0 \text{ cm}.$$

The negative sign tells us the image is virtual.

$$|M| = \left|\frac{s_i}{s_0}\right| = \left|\frac{-24.0 \text{ cm}}{8.0 \text{ cm}}\right| = 3.0.$$

Note: The real image of part (a) is inverted while the virtual image of part (b) is upright or erect. To build this result into the equations, define the magnification, not its magnitude with a negative sign, that is,

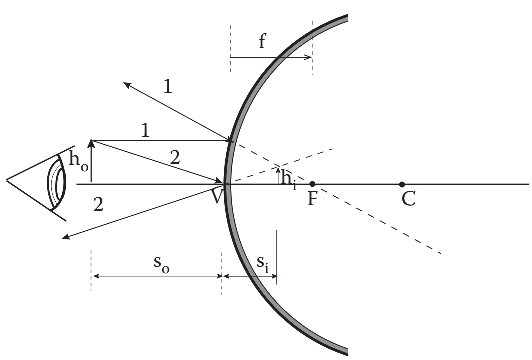
$$M \equiv \frac{-s_i}{s_0} = \frac{h_i}{h_0}. \tag{20.10}$$

Thus, if M is positive, the image is upright, if negative the image is inverted. The mirror equation 20.9, although derived for a concave mirror, also applies to convex mirrors (see the figure below) and will yield correct results if a consistent sign convention is adopted.

In this chapter, the following *sign convention* will be adopted:

1. The focal length (f) is positive for a converging (concave) mirror and negative for a diverging (convex) mirror.

   Note: Another way of perceiving this rule is if the center of curvature (C) is on the reflecting side of the mirror, (f) is positive, if not, (f) is negative.



Images formed by a convex mirror are always virtual.

2. The object distance (s_0) is positive if the object is located on the reflecting side of the mirror. That is, (s_0) is positive for a real object.
3. The image distance (s_i) is positive if light, *after reflection*, travels from the reflecting surface to the image. Otherwise it is negative.
4. The object and image heights h_0 and h_i are positive if upright.

These rules may appear awkward as stated, but their utility will become evident with use. In particular, (a) they will apply, with minimum modification, to thin lens systems and (b) to combinations of mirrors where the image of one mirror is treated as the object of the second mirror.

EXAMPLE 20.2

For the mirror in the figure above, determine the location of the image and its magnification for an object if $h_0 = 6.0$ cm, $f = -12.0$ cm and (a) $s_0 = 18.0$ cm and (b) $s_0 = 8.0$ cm.

Solution

a. Equation 20.9 gives

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_0} = \frac{1}{(-12.0\text{ cm})} - \frac{1}{(18.0\text{ cm})} = \frac{-3.0 - 2.0}{36.0\text{ cm}} = \frac{-5.0}{36.0\text{ cm}},$$

$$s_i = -\frac{36.0\text{ cm}}{5.0} = -7.20\text{ cm}.$$

The negative sign, by rule 3, implies that the image is formed *behind* the reflecting surface. Since physical light rays do NOT pass through this image plane, the image is called a *virtual image*. By Equation 20.10,

$$M = \frac{(-s_i)}{s_0} = -\frac{(-7.2\text{ cm})}{18.0\text{ cm}} = +0.400 = \frac{h_i}{h_0}.$$

The positive sign indicates that the image is upright. Also, it is smaller than the object, that is,

$$hi = M(h0) = +0.400(6.0\text{ cm}) = 2.4\text{ cm}.$$

b. Again, by Equation 20.9,

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_0} = \frac{1}{(-12.0\text{ cm})} - \frac{1}{(8.0\text{ cm})} = \frac{-2 - 3}{(24.0\text{ cm})} = \frac{-5}{(24.0\text{ cm})},$$

$$s_i = \frac{24.0\text{ cm}}{-5} = -4.8\text{ cm}.$$

The negative sign implies a virtual image at 4.8 cm behind the reflecting surface. Also,

$$M = \frac{(-s_i)}{(s_0)} = \frac{-(-4.8\text{ cm})}{8.0\text{ cm}} = +0.60 = \frac{h_i}{h_0}$$

or

$$h_i = M(h_0) = (+0.60)(6.0\text{ cm}) = 3.6\text{ cm}.$$

20.4 REFRACTION

The velocity of light in a transparent or translucent material medium (v) is less than its velocity in air or in vacuum (c). This is so because light transmission through a material medium is partially an absorption–reemission process. That is, the atoms or molecules in the medium absorb many of the light particles (photons), hold them for microseconds or less, and reemit them to form wavefronts that propagate through the material medium. The timing of the reemission process depends on the specific microstructure of the material, so light has a different velocity in different materials. This property is characterized by a quantity called the *index of refraction* (n), which is defined as

$$n \equiv \frac{c}{v_{med}}. \tag{20.11}$$

Since $v_{med} \leq c$, $n \geq 1.0$.

EXAMPLE 20.3

The indices of refraction of pure water and diamond are 1.333 and 2.417, respectively. Determine the value of the velocity of light in these materials.

Solution

Using Equation 20.11,

$$v_{med} = v_{water} = \frac{c}{n_{water}} = \frac{3.00 \times 10^8 \text{ m/s}}{1.333} = 2.25 \times 10^8 \text{ m/s}.$$

Note that

$$v_{water} = \frac{c}{n} = 1.333 \text{ implies } n = \frac{c}{(4/3)c} = \frac{3}{4}c = 0.75 \text{ c}.$$

For diamond

$$v_{dia} = \frac{c}{n} = \frac{3.00 \times 10^8 \text{ m/s}}{2.417} = 1.24 \times 10^8 \text{ m/s} \cong 0.41c.$$

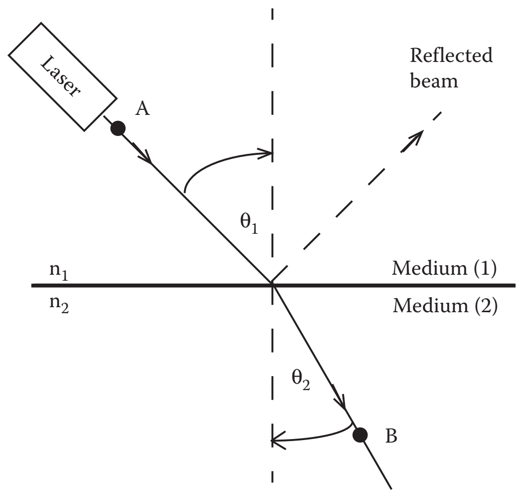
Law of Refraction (Snell's Law)

Consider a beam of light originating at point A in medium 1 and propagating to and through point B in medium 2 (the following figure). The beam in medium 2 is the refracted beam. If the interface between the two media is smooth, a portion of the incident beam will also be specularly reflected and obey the law of reflection. These reflected beams are generally omitted in refraction problems. Suppose the location of point B and A are initially assumed to be arbitrary and an equation for the path length between points A and B, in terms of light velocity and time, is constructed. Recall that Fermat's principle implies that the light will follow that path whereby the *time* to go from A to B is a *minimum*. If the constructed equation is minimized with respect to time, *Snell's law of refraction* results. That is,

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \text{ (Snell's law of refraction).} \tag{20.12}$$

EXAMPLE 20.4

In the following figure, let medium 1 be air with $n_1 = 1.000$ and medium 2 water with $n_2 = 1.333$. If the incident angle is $\theta_1 = 45.0°$, determine the angle of refraction θ_2.



Light rays are refracted (bent) at the interface between media of different indices of refraction.

Solution

Using Equation 20.12,

$$\sin\theta_2 = \left(\frac{n_1}{n_2}\right)\sin\theta_1 = \left(\frac{1.000}{1.333}\right)\sin 45.0° = (0.750)(0.707) = 0.530$$

or

$$\theta_2 = \sin^{-1}(0.530) = 32.0°.$$

Note: Snell's law is reflexive, that is, also works in reverse. If the light source is in medium 2 and the beam direction is reversed so that θ_2 is the incident angle, then θ_1 will be the angle of refraction.

EXAMPLE 20.5

In the figure above, let medium 1 be water and medium 2 a thick slab of crown glass with $n_2 = 1.523$ and laying on the bottom of swimming pool. A boy, in the water, with a water proof laser pointer directs its beam at the glass at an incident angle $\theta_1 = 45.0°$ with respect to the normal. Determine the angle of refraction.

Solution

Again, Equation 20.12 gives

$$\sin\theta_2 = \left(\frac{n_1}{n_2}\right)\sin\theta_2 = \left(\frac{1.333}{1.523}\right)\sin(45.0°) = 0.619,$$

so

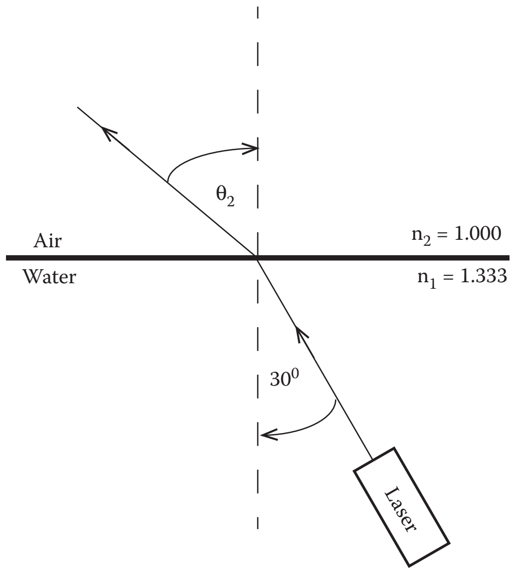
$$\theta_2 = \sin^{-1}(0.619) = 38.2°.$$

Total Internal Reflection

In Examples 20.4 and 20.5, the light source was located in the lower index of refraction medium and the light was directed into the higher index medium. Consider the reverse situation where the beam is directed from high to low index as illustrated in the next example.

EXAMPLE 20.6

A girl, under water in a swimming pool, directs the light from a waterproof laser pointer upward at an angle of $\theta_1 = 30.0°$ (the figure below). Determine the refraction angle of the beam.



Refracted light rays are reflexive. The ray diagram is valid also when ray directions are reversed.

Solution

Snell's law gives

$$\sin\theta_2 = \left(\frac{n_1}{n_2}\right)\sin\theta_1 = \left(\frac{1.333}{1.000}\right)\sin(30.0°) = 0.667$$

or

$$\theta_2 = \sin^{-1}(0.667) = 41.8°.$$

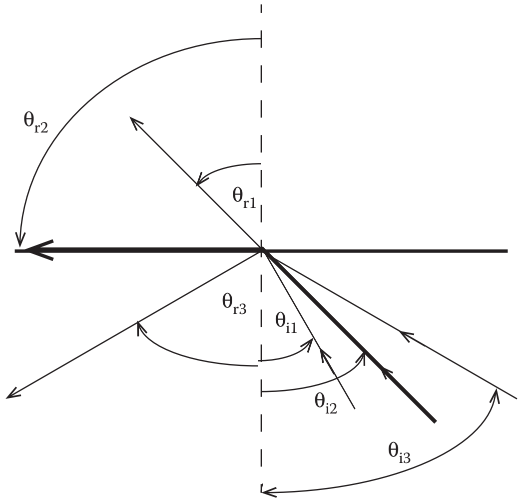
For this situation, $\theta_2 > \theta_1$. If θ_1 is increased through various values θ_{i1}, θ_{i2}, θ_{i3}, and so on, θ_2 will correspondingly increase to the values θ_{r1}, θ_{r2}, θ_{r3}, and so on. Here, the subscript "i" represents the "incident" beam and "r" the "refracted" beam.

The following figure indicates that, at some particular value of θ_i, call it the *critical angle* θ_c, the refraction angle θ_r will be 90°. That is, the refracted beam will propagate along the interface between the two media. To calculate the critical angle, insert $\theta_r = 90.0°$ into Snell's law and solve for the incident angle. Thus,

$$n_i \sin\theta_c = n_r \sin\theta_r$$

or

$$\sin\theta_c = \frac{n_r \sin(90.0°)}{n_i} = \frac{n_r}{n_i}$$



When the critical angle is exceeded, total internal reflection occurs.

so

$$\theta_c = \sin^{-1}\left(\frac{n_r}{n_i}\right)\text{(critical angle).} \tag{20.13}$$

EXAMPLE 20.7

Determine the critical angle for the air–water interface in Example 20.6.

Solution

Direct application of Equation 20.13 gives

$$\theta_c = \sin^{-1}\left(\frac{n_r}{n_i}\right) = \sin^{-1}\left(\frac{1.000}{1.333}\right) = 48.6°.$$

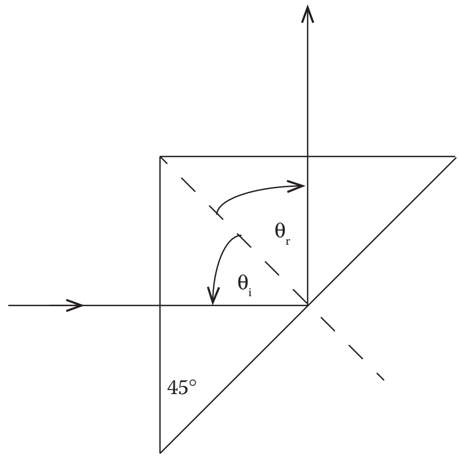
Note: Since Snell's law is reflexive, any object located infinitesimally above the surface can be seen by the submerged girl if she looks up toward the surface at an angle of 48.6°. In reality, waves on the unsmooth water surface would distort the image.

If the critical angle is exceeded, as illustrated by angle θ_{i3} in the figure above, the incident beam will be *reflected*, not refracted and none of the light will be transmitted through the interface. This is called *total internal reflection*. The reflection is indeed *total* if the interface is very smooth. That is, no EMW energy leaves the medium in which the incident beam originated. Total internal reflection has many practical uses. Two that are addressed here are (a) reflecting glass prisms used in periscopes and binoculars and (b) fiber optic cables.

a. The crown glass prism in the following figure has an index of refraction $n_1 = 1.523$ and is situated in air with $n_2 = 1.000$. The beam of light enters from the left at an angle of 90.0°, that is, along the normal to the interface, and so is not deviated from its original straight-line path.

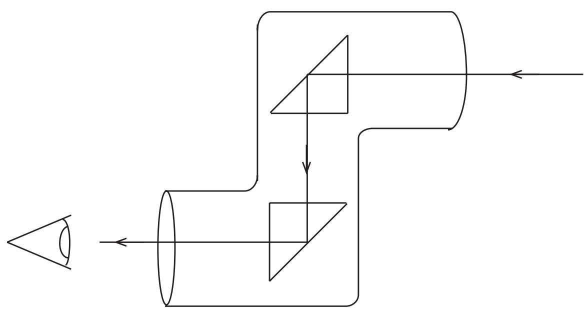
   The prism is positioned so that the beam inside strikes the second interface at an angle $\theta_i = 45.0°$. The critical angle for the prism material is

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) = \sin^{-1}\left(\frac{1.000}{1.532}\right) = 41.0°,$$



Total internal reflection in a glass prism.

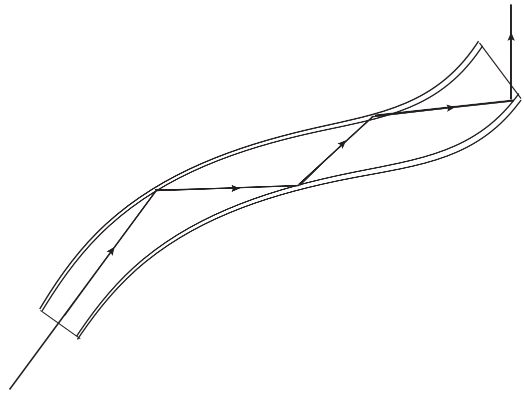
So, because $\theta_i > \theta_c$, total internal reflection occurs. For the reflection, $\theta_r = \theta_i$, so the beam is reflected at an angle $\theta_r = 45.0°$ relative to the normal. Clearly, a plane mirror could be used to similarly redirect the incident beam, but the reflection would not be total; that is, some of the incident light energy would be absorbed by the mirror. For conditions of low light intensity or for multiple reflections, total internal reflection is better. Two such prisms combined in a closed tube can be used to construct an elementary periscope, as in the figure below.



Total internal reflection used to redirect light rays in a periscope.

b. Another important application of total internal reflection is in the construction of "light pipes" or fiber optics cables where plane mirrors are not a viable alternative. The light pipe consists of a long thin cylindrical thread of transparent glass or plastic fiber, with relatively high index of refraction (see the following figure). The fiber is enclosed in a concentric cylindrical coating of thin transparent glass with low index of refraction. The coating, called the *cladding*, increases the transmission efficiency of the fiber. Light enters the fiber at one end, encounters the interface with the cladding at an incident angle that is greater than the critical angle, and so is totally internally reflected. The reflected light continues through the fiber until its next encounter with the interface. Again, total internal reflection occurs.

This zigzag process continues and so the light is confined to the interior of the fiber and is propagated through it, much like flowing water is confined to the interior of a pipe. If the glass fiber has few impurities, the light can travel long distances (many miles) before its intensity is appreciably attenuated. In practice, many fibers are bundled together to create cables. It is possible to send much more information through these cables than through electrical cables of the same size.



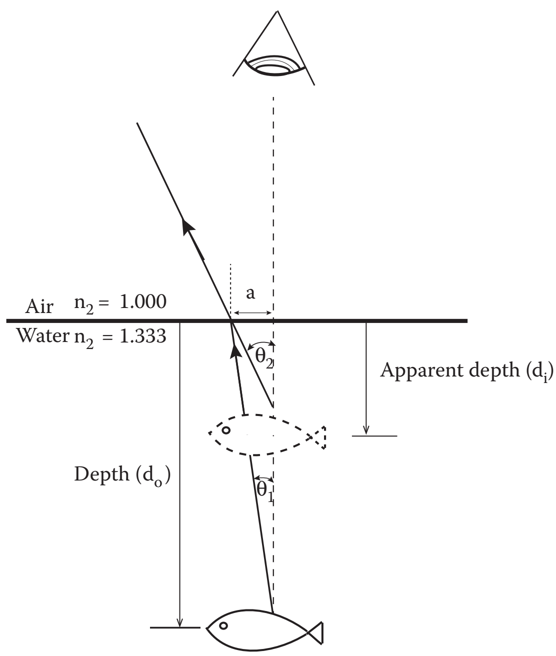
Internal reflections in a fiber optics cable or "light pipe."

Apparent Depth

Another consequence of Snell's law is that objects submerged in a transparent liquid appear closer to the interface (surface) than they actually are, according to an observer positioned in air.

The figure below shows a fish, submerged at a depth d_0. Light from above reflects off the fish and propagates upward to the surface where it is refracted. One such ray, reflected from the center of the fish, makes an angle θ_1 with the normal and is refracted through the angle θ_2. For an observer positioned above the fish, the angles θ_1 and θ_2 are small (≤15°), so the approximation $\sin\theta \cong \tan\theta$ is reasonable. Snell's law can be written as

$$n_1 \tan\theta_1 \cong n_2 \tan\theta_2$$



Objects submerged in water appear closer to the surface than their true depth.

or

$$n_1\left(\frac{a}{d_0}\right) \cong n_2\left(\frac{a}{d_i}\right).$$

So, the apparent depth is

$$d_i = d_0\left(\frac{n_2}{n_1}\right). \tag{20.14}$$

EXAMPLE 20.8

The fish in the figure above is swimming at 1.20 m below the lake's surface. How far below the surface does it appear to be an observer in a fishing boat directly above the fish?

Solution

Direct substitution into Equation 20.14 gives

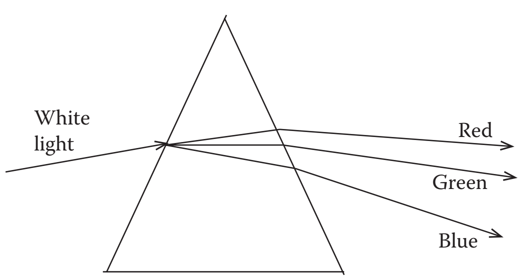
$$d_i = d_0\left(\frac{n_2}{n_1}\right) = (1.20\ \text{m})\left(\frac{1.000}{1.333}\right) = 0.900\ \text{m} = 90.0\ \text{cm}.$$

Dispersion

It is found experimentally that the velocity of light in a transparent material medium depends on the frequency of the light. Higher frequency rays travel more slowly than those of lower frequency. Thus, $v \approx 1/f$.

Since the index of refraction is inversely proportional to the velocity, then $n \approx 1/v$. So, $n \propto 1/(1/f) \propto f$; that is, n depends directly on the frequency or, correspondingly, inversely on the wavelength.

Violet light, with the highest frequency in the visible spectrum, has the largest n and will be bent (refracted) the most at the interface between two transparent media. Red light will have the smallest n and be refracted least. This effect is illustrated in the figure below that shows white light incident on a glass prism. The white light is refracted at the first interface and decomposed into its constituent colors, that is, dispersed. It travels through the prism and is refracted again at the second interface. This decomposition into constituent colors is called *dispersion* and is the basis of effects such as the rainbow.



The index of refraction depends on the inverse of the wavelength of the light.

20.5 THIN LENSES

A *lens* is a piece of transparent material with curved surfaces, constructed to control the refraction of light. Usually, the index of refraction of the lens material (n_ℓ) has a value >1.00. The cross section, through the diameter, of a circular converging lens is shown in Figure 20.10.

It is constructed of spherical surfaces whose radii of curvature are R_1 and R_2, respectively. These radii need not be equal, and indeed one may be equal to infinity so that its surface is a flat plane. The lens is immersed in a transparent medium whose index of refraction is n_m. Typically, but not always, $n_m = 1.00$; that is, the medium is air. Note that the normal to surface "1" is along radius R_1 and to surface "2" along R_2. Figure 20.11 shows a ray striking surface 1, being refracted toward its normal inside the lens, striking surface 2 and being refracted away from its normal outside that surface. The path of the ray indicates that the lens makes it converge toward the principal axis of the lens.

A *thin lens* is one whose thickness (d) is much less than the length C_1 to C_2. The lens of Figure 20.11 has a cross section that is thicker at its center and thinner away from its center, that is, away from its principal axis. Such a lens is called a *double-convex lens*. If the reverse is true, that is, a cross section thinner at the center and thicker at its periphery, it is called a *double-concave lens*. If rays of light originate in a source that is far from the lens, they will be approximately parallel to the principal axis. They are called *paraxial rays*. Ray tracing, employing Snell's law, shows that paraxial rays, after passing through the lens, will intersect at a common point called the *focal point* (Figure 20.12).

The focal point (F) is located at a distance (f) from the center of the lens and is called the *focal length*. An observer, situated to the right of (F), would interpret the rays as originating at the focal point. Note that although the rays are refracted at *both* interfaces, it is conventional and convenient to draw just one "bending" at the center of the cross section. Also, if the paraxial rays in Figure 20.12 had originated

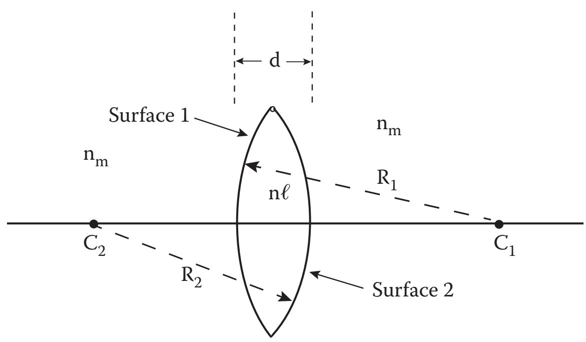


FIGURE 20.10 A converging lens viewed in cross section.

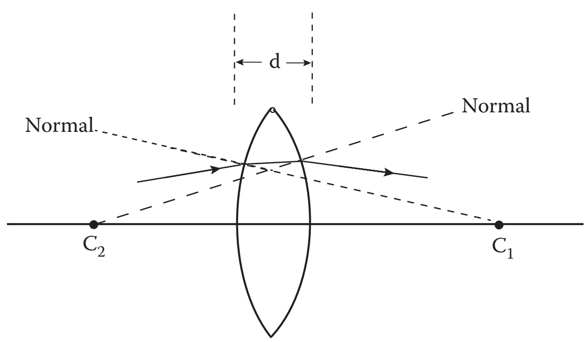


FIGURE 20.11 A lens refracts light rays at both interfaces.

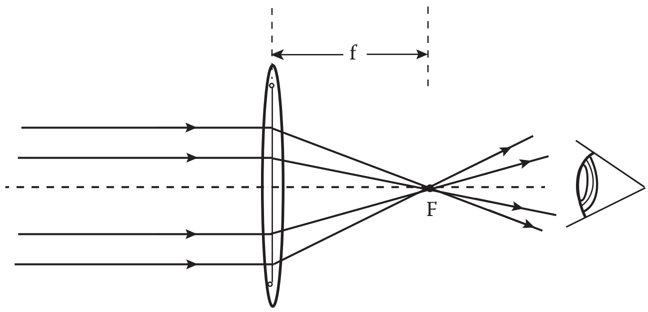


FIGURE 20.12 All paraxial rays pass through the focal point of the lens after refraction.

to the right of the lens, the focal (F) would occur to the left of the lens. So, each thin lens has two focal points, situated at equal distances (f) on either side of the lens. It can be shown that the focal length is determined by the radii R_1 and R_2 and the index of refraction of the lens material. The relation is

$$\frac{1}{f} = \frac{(n_\ell - n_m)}{n_m}\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \tag{20.15}$$

EXAMPLE 20.9

A plano-convex lens is made of flint glass which has an index of refraction of $n_\ell = 1.66$. Its radii of curvature are $R_1 = +15.0$ cm and $R_2 = -\infty$. Determine its focal length in (a) air, $n_m = 1.00$, and (b) water, $n_m = 1.33$.

SOLUTION

a. Using Equation 20.15,

$$\frac{1}{f} = \frac{(n_\ell - n_m)}{n_m}\left[\frac{1}{R_1} - \frac{1}{R_2}\right] = \frac{(1.66 - 1.00)}{1.00}\left[\frac{1}{(15.0\text{ cm})} - \frac{1}{(-\infty)}\right]$$
$$= (0.66)\left[(0.067) - (0.00)\right] = 4.4 \times 10^{-2}\text{ cm}^{-1}$$

or

$$f = 23\text{ cm}.$$

b.

$$\frac{1}{f} = \frac{(1.66 - 1.33)}{1.33}\left[\frac{1}{(915.0\text{ cm})} - \frac{1}{(-\infty)}\right] = (0.248)(0.067) = 1.7 \times 10^{-2}\text{ cm}^{-1}$$

or

$$f = 60\text{ cm}.$$

Additionally, light rays are reflexive, that is, their ray diagrams are also valid if the directions of the arrows on the rays are reversed. This property suggests that if the source of light was placed at the focal point (F), the rays, after refraction, would be paraxial. This is a way to create paraxial rays, namely place the light source at the focal point of a converging lens. This property is exploited in both the headlights and tail lights of automobiles.

Figure 20.13 shows paraxial rays refracted by a double-concave lens. The rays diverge after refraction, but if extended backward, appear to an observer to originate from a point, which is the focal point. Note that for paraxial rays, real light energy passes through the focal point of a convex lens but not through (F) of a concave lens.

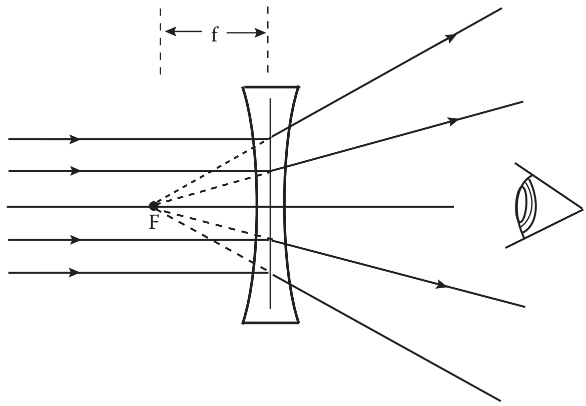


FIGURE 20.13 Paraxial rays appear to originate and converge from the focal point of a concave lens, when viewed from downstream.

If the source of light (the object) is close, not far from the lens, an image is formed (Figure 20.14). As with mirrors, the location of the image, formed by lenses, can be found either graphically or analytically.

Graphically, the properties of several select rays allow the determination of image location and size. In Figure 20.14, ray one is paraxial and passes through (F) after refraction. Ray two passes through the center of the lens at the principal axis and is essentially undeflected for a thin lens. Ray three is chosen to pass through the focal point before refraction. After refraction, ray three is paraxial. These three specifically chosen rays intersect in a plane, the *image plane*, and form an image of the object. If the diagram is drawn to scale, the image distance (s_i) and size (h_i) can be measured off the diagram.

Analytically, the image location and size can be determined from an equation relating (s_i) to (s_0) and (f). To derive the equation, consider the parameters in Figure 20.14 both before and after refraction.

$$\tan\alpha = \frac{h_0}{s_0} = \frac{|h_i|}{s_i}$$

or

$$\frac{|h_i|}{h_0} = \frac{s_i}{s_0}. \tag{20.16}$$

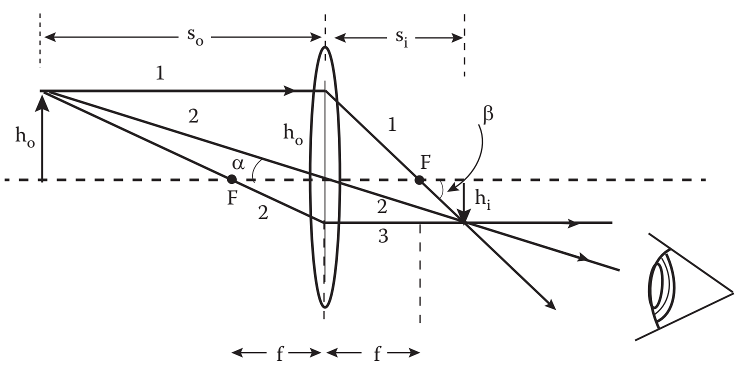


FIGURE 20.14 The "mapping" of light rays by a convex lens to form an image.

And

$$\tan\beta = \frac{h_0}{f} = \frac{|h_i|}{(s_i - f)},$$

so,

$$\frac{|h_i|}{h_0} = \frac{(s_i - f)}{f}. \quad (20.17)$$

Equating the right sides of Equations 20.16 and 20.17 and rearranging gives

$$\frac{1}{s_0} + \frac{1}{s_i} = \frac{1}{f}. \quad (20.18)$$

Equation 20.18 is the same as Equation 20.9, the mirror equation, and may be called the thin lens equation. If combined with Equation 20.15, the thin lens equation becomes

$$\frac{1}{s_0} + \frac{1}{s_i} = \frac{1}{f} = \frac{(n_\ell - n_m)}{n_m}\left[\frac{1}{R_1} - \frac{1}{R_2}\right]. \quad (20.19)$$

Note: Although Equation 20.19 was derived for a convex lens, it can be shown to be valid for concave lenses also. Figure 20.15 is a diagram of some specially chosen rays for a concave lens.

Ray one is paraxial and diverges after refraction. Its backward extension passes through the focal point (F). Ray two passes through the lens center and is undeflected by a thin lens. The two rays intersect in the image plane and contribute to forming the image. Note that if a small screen was placed in the image plane (small, so as not to block all of the light from the object), no image would be seen on it. But to an observer to the right of the lens, the light would appear to come from the image plane. Such an image is called a *virtual image*.

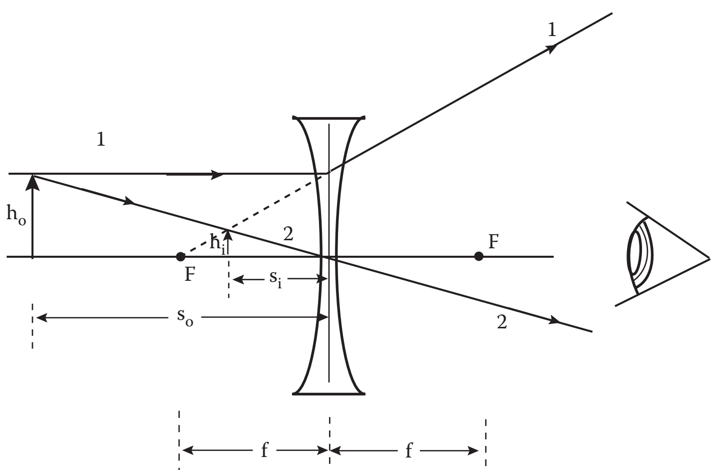


FIGURE 20.15 The "mapping" of light rays by a concave lens to form an image.

20.5.1 Sign Convention

To use Equation 20.19 correctly, a sign convention is needed. The following sign convention will be used:

1. The focal length (f) is positive for a convex lens and negative for a concave lens.
2. The object distance (s_0) is positive for a real object.
3. The image distance (s_i) is positive if light, after refraction, travels from the lens to the image. Otherwise, it is negative.
4. The object and image heights are positive if upright.

EXAMPLE 20.10

Determine the location (s_i) and type of the image formed when an object of height $h_0 = 4.0$ cm is placed (a) 18.0 cm and (b) 9.0 cm from a convex lens of focal length f = 12.0 cm.

Solution

a. Using Equation 20.19,

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_0} = \frac{1}{(+12.0 \text{ cm})} - \frac{1}{(18.0 \text{ cm})} = \frac{3-2}{36.0 \text{ cm}}$$

or

$$s_i = \frac{36.0 \text{ cm}}{1.0} = 36.0 \text{ cm}.$$

Since s_i is positive, the image is real.

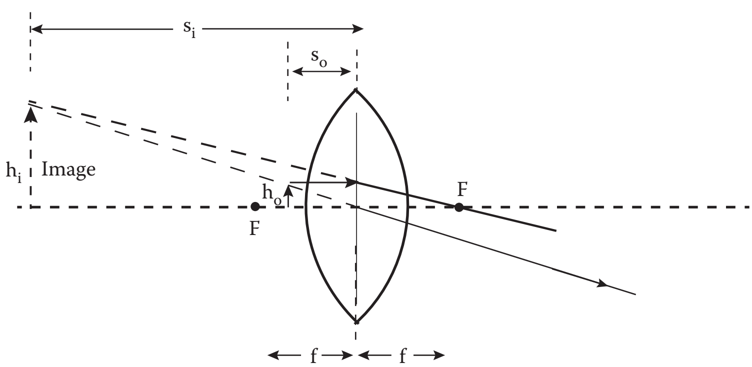
b.

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_0} = \frac{1}{(+12.0 \text{ cm})} - \frac{1}{(9.0 \text{ cm})} = \frac{3-4}{36.0 \text{ cm}}$$

or

$$s_i = \frac{36.0 \text{ cm}}{(-1.0)} = -36.0 \text{ cm}.$$

A negative image distance means the image is virtual. A virtual image cannot be projected on a screen since no EMW energy is propagated into the image plane after refraction. The figure below illustrates the solution to part (b) of this example.



If the object is inside the focal length of a convex lens, the image is virtual.

EXAMPLE 20.11

Determine the location and type of image formed when an object of height $h_0 = 4.0$ cm is placed (a) 18.0 cm, and (b) 9.0 cm from a concave lens of focal length $f = -12.0$ cm. *Note:* This example is similar to Example 20.10 but replaces the convex lens with a concave lens.

Solution

a. Equation 20.19 gives

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_0} = \frac{1}{(-12.0\text{ cm})} - \frac{1}{(18.0\text{ cm})} = \frac{(-3-2)}{(36.0\text{ cm})} = \frac{(-5)}{(36.0\text{ cm})}$$

or

$$s_i = \frac{(36.0\text{ cm})}{(-5)} = -7.2\text{ cm.}$$

The negative sign implies a virtual image.

b.

$$\frac{1}{s_i} = \frac{1}{(-12.0\text{ cm})} - \frac{1}{(9.0\text{ cm})} = \frac{(-3-4)}{(36.0\text{ cm})} = \frac{(-7)}{(36.0\text{ cm})}$$

or

$$s_i = -5.1\text{ cm.}$$

Again, a virtual image.

The images formed by thin lenses, as is the case with mirrors, can be either upright or inverted. The lateral *magnification* (M) is useful to determine image orientation analytically. Recall that $M \equiv (|h_i|/h_0)$. In the second figure under Example 20.9, which was used to derive the thin lens equation, the image is inverted, but that inversion was ignored in Equation 20.16. To account for inversions, write Equation 20.16 as

$$\frac{h_i}{h_0} = \frac{-s_i}{s_0} \equiv M. \qquad (20.20)$$

Thus, if the calculation of M yields a negative number, the image is inverted relative to the object; that is, the image and object have opposite orientations.

EXAMPLE 20.12

Use Equation 20.20 to determine the size and orientation of the images calculated in Examples 20.9 and 20.10.

Solution

In Example 20.10a, $h_0 = 4.0$ cm, $s_0 = 18.0$ cm, and $s_i = 36.0$ cm, so

$$M = \frac{-s_i}{s_0} = \frac{-36.0\text{ cm}}{18.0\text{ cm}} = -2 = \frac{h_i}{h_0}$$

and

$$h_i = (-2)\,h_0 = (-2)(4.0\text{ cm}) = -8.0\text{ cm.}$$

The image is twice as large as the object and inverted.

In Equation 20.10b, $s_0 = 9.0$ cm and $s_0 = -36.0$ cm, then

$$M = \frac{-s_i}{s_0} = \frac{-(-36.0 \text{ cm})}{9.0 \text{ cm}} = +4.0.$$

The image is four times larger than the object and is upright.

In Example 20.11a, $h_0 = 4.0$ cm, $s_0 = 18.0$ cm, and $s_i = -7.2$ cm, so

$$M = \frac{-s_i}{s_0} = \frac{-(-7.2 \text{ cm})}{18.0 \text{ cm}} = +0.40.$$

So, the image is upright but smaller than the object. Thus,

$$M = \frac{h_i}{h_0} = +0.40.$$

So,

$$h_i = (0.40)h_0 = (0.40)(4.0 \text{ cm}) = 1.6 \text{ cm}.$$

In Example 20.11b, $s_0 = 9.0$ cm and $s_i = -5.1$ cm, then

$$M = \frac{-s_i}{s_0} = \frac{-(-5.1 \text{ cm})}{9.0 \text{ cm}} = +0.57 = \frac{h_i}{h_0}$$

or

$$h_i = (+0.57)(h_0) = (0.57)(4.0 \text{ cm}) = 2.3 \text{ cm}.$$

The image is upright and smaller.

Optical instruments such as the microscope and telescope exploit the properties of two or more lenses in combination. The final image location, size and orientation again, can be found either graphically or analytically. The graphical method is tedious and requires a knowledge of the radii of curvature (so as to determine the normal to the surfaces) of each interface. When this method is used, a simple generalization can be made. It is (a) the thin lens equation may be applied to each lens separately and (b) the image of the first lens can be treated as the object of the second. Once the location of this first image has been determined, the first lens can be ignored and the second lens can be treated separately. Then, the image of the second lens is treated as the object of the third, and so on.

This generalization permits the use of the simpler analytical method instead of the more tedious graphical method.

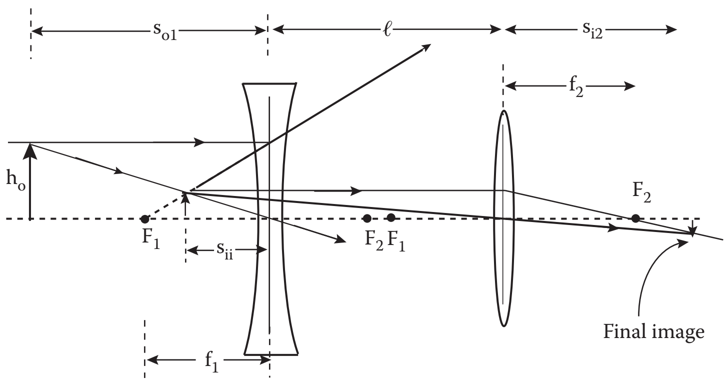
EXAMPLE 20.13

The following figure illustrates a two-lens system with an object to the left of the concave lens. Numerically, $f_1 = -12.0$ cm, $f_2 = +9.0$ cm, $s_{01} = 24.0$ cm, and $\ell = 28.0$ cm.

Solution

Applying the thin lens equation to the first lens gives

$$\frac{1}{s_{i1}} = \frac{1}{f_1} - \frac{1}{s_{01}} = \frac{1}{(-12.0 \text{ cm})} - \frac{1}{(24.0 \text{ cm})} = \frac{(-2-1)}{(24.0 \text{ cm})}$$



A two-lens system. The image of the first lens should be treated as the object of the second lens.

or

$$s_{i1} = \frac{(24.0\text{ cm})}{(-3)} = -8.0\text{ cm}.$$

The negative sign indicates that this first image is *virtual* and is to the left of the concave lens. Its magnification is

$$M = \frac{-s_i}{s_0} = \frac{-(-8.0\text{ cm})}{24.0\text{ cm}} = +\frac{1}{3} = 0.33.$$

The first image is only one-third as large as the object. The positive sign indicates that the image is upright; that is, it has the same vertical orientation as the object. This image is treated as the object of the second lens and the presence of the first lens is now ignored. The object distance is

$$s_{02} = (s_{i1} + \ell) = (8.0\text{ cm} + 28.0\text{ cm}) = 36.0\text{ cm},$$

so,

$$\frac{1}{s_{i2}} = \frac{1}{f_2} - \frac{1}{s_{02}} = \frac{1}{(9.0\text{ cm})} - \frac{1}{(36.0\text{ cm})} = \frac{(4-1)}{(36.0\text{ cm})} = \frac{(3)}{(36.0\text{ cm})}$$

$$s_{i2} = 12.0\text{ cm}.$$

Since it is positive, (s_{i2}) is to the right of lens two and the image is *real*. The magnification, due to lens two only, is

$$M_2 = \frac{-s_{i2}}{s_{02}} = \frac{-12.0\text{ cm}}{36.0\text{ cm}} = -\frac{1}{3} = -0.33.$$

Lens two inverts its object. The total magnification of the system is

$$M_{total} = M_1 M_2 = \left(\frac{-s_{i1}}{s_{01}}\right)\left(\frac{-s_{i2}}{s_{02}}\right) = \left[\frac{-(-8.0\text{ cm})}{24.0\text{ cm}}\right]\left(\frac{-12.0\text{ cm}}{36.0\text{ cm}}\right) = \left(\frac{1}{3}\right)\left(\frac{-1}{3}\right) = -0.11.$$

The final image is real, inverted, and only one-ninth as large as the object.

EXAMPLE 20.14

Suppose the lenses in Example 20.13 are interchanged, but all other parameters remain the same. Determine the location and nature of the final image.

Solution

The subscripts "1" and "2" now refer to the convex and concave lenses, respectively. For the first lens

$$\frac{1}{s_{i1}} = \frac{1}{f_1} - \frac{1}{s_{01}} = \frac{1}{(+9.0\text{ cm})} - \frac{1}{(24.0\text{ cm})} = \frac{(24.0\text{ cm}) - (9.0\text{ cm})}{(9.0\text{ cm})(24.0\text{ cm})} = \frac{15.0\text{ cm}}{216\text{ cm}^2}$$

or

$$s_{i1} = \frac{216\text{ cm}^2}{15.0\text{ cm}} = 14.4\text{ cm},$$

which is to the right of lens one.

Then, $s_{02} = (\ell - s_{i1}) = (28.0\text{ cm} - 14.4\text{ cm}) = 13.6\text{ cm}$. So,

$$\frac{1}{s_{i2}} = \frac{1}{f_2} - \frac{1}{s_{02}} = \frac{1}{(-12.0\text{ cm})} - \frac{1}{(13.6\text{ cm})} = \frac{(-13.6\text{ cm} - 12.0\text{ cm})}{163\text{ cm}^2} = \frac{(-25.6\text{ cm})}{(163\text{ cm}^2)}$$

or

$$s_{i2} = -6.37\text{ cm}.$$

The negative sign implies that the image is virtual and to the left of the second (concave) lens. The total magnification is

$$M_{total} = \left(\frac{-s_{i1}}{s_{01}}\right)\left(\frac{-s_{i2}}{s_{02}}\right) = \left(\frac{-14.4\text{ cm}}{24.0\text{ cm}}\right)\left[\frac{-(-6.37\text{ cm})}{13.6\text{ cm}}\right]$$

or

$$M = (-0.0600)(0.468) = -0.281.$$

The final image is virtual, inverted, and smaller relative to the object.

EXAMPLE 20.15: VIRTUAL OBJECT

The following figure shows a two-lens system with an object at a distance of $s_{01} = 18.0$ cm to the left of the first lens. Additionally, $f_1 = 12.0$ cm, $f_2 = 20.0$ cm, and the separation is $\ell = 24.0$ cm. Determine the location and nature of the final image.

Solution

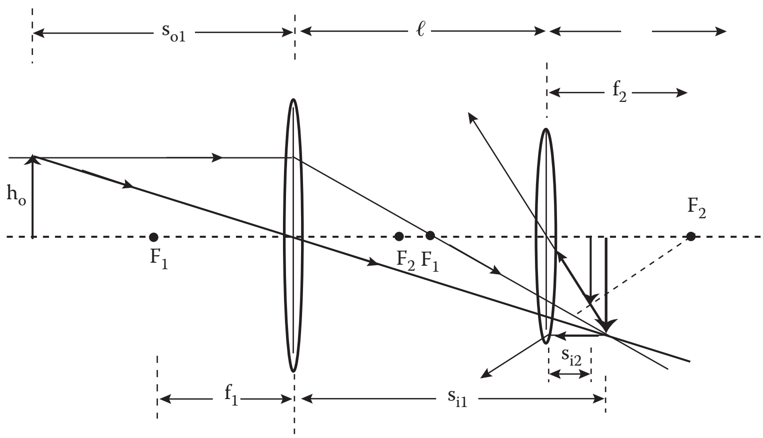
Treating the first lens separately,

$$\frac{1}{s_{i1}} = \frac{1}{f_1} - \frac{1}{s_{01}} = \frac{1}{(+12.0\text{ cm})} - \frac{1}{(18.0\text{ cm})} = \frac{(3-2)}{36.0\text{ cm}} = \frac{1}{36.0\text{ cm}}$$

or

$$s_{i1} = 36.0\text{ cm}$$

to the right of the first lens *and* also to the right of the second lens.



An example of a virtual object.

After passing through the first lens, the light rays are intercepted by the second lens. This refraction is ignored at this stage of the calculation. From the figure above, it is seen that the rays from the first image do not travel *backward* through the second lens. Thus, the image of lens one is the *virtual object* of lens two. So,

$$s_{02} = (\ell - s_{i1}) = (24.0 \text{ cm} - 36.0 \text{ cm}) = -12.0 \text{ cm},$$

then

$$\frac{1}{s_{i2}} = \frac{1}{f_2} - \frac{1}{s_{02}} = \frac{1}{(+20.0 \text{ cm})} - \frac{1}{(-12.0 \text{ cm})} = \frac{(3+5)}{(60.0 \text{ cm})} = \frac{8}{60.0 \text{ cm}}$$

$$s_{i1} = 7.5 \text{ cm}.$$

The final image is real and located 7.5 cm to the right of lens two. The total magnification for the system is

$$M_T = M_1 M_2 = \left(\frac{-s_{i1}}{s_{01}}\right)\left(\frac{-s_{i2}}{s_{02}}\right) = \left(\frac{-36.0 \text{ cm}}{18.0 \text{ cm}}\right)\left(\frac{-7.5 \text{ cm}}{-12.0 \text{ cm}}\right)$$

$$M_T = -1.3.$$

The final image is inverted and slightly larger than the original object. In this example, if the second lens was a diverging lens (concave) instead of convex, would the final image be real or virtual? To check, suppose $f_2 = -20$ cm. The first stage of this calculation is the same as the first stage of the previous solution, so

$$s_{i1} = 36.0 \text{ cm} \quad \text{and} \quad s_{02} = -12.0 \text{ cm}.$$

Then,

$$\frac{1}{s_{i2}} = \frac{i}{f_2} - \frac{1}{s_{02}} = \frac{1}{(-20.0 \text{ cm})} - \frac{1}{(-12.0 \text{ cm})} = \frac{(-3+5)}{(60.0 \text{ cm})} = \frac{2}{(60.0 \text{ cm})}$$

$$s_{i2} = 30.0 \text{ cm}.$$

So the image is real. Note that even though the second lens is concave, a real image is produced for a virtual object.

$$M_T = M_1M_2 = \left(\frac{-s_{i1}}{s_{01}}\right)\left(\frac{-s_{i2}}{s_{02}}\right) = \left(\frac{-36.0\text{ cm}}{18.0\text{ cm}}\right)\left(\frac{-30.0\text{ cm}}{-12.0\text{ cm}}\right) = -5.0.$$

Also, the final image is inverted and five times larger than the object.

20.6 OPTICAL INSTRUMENTS

Optical instruments such as binoculars, refracting telescopes, and microscopes use two or more lenses in combination. In the simplest instruments, made with just two lenses, both are convex, that is, converging. Both are placed inside a rigid tube or tubes that can be brought up to the eye for viewing. The lens closest to the eye appropriately is called the *eyepiece* or *ocular*. The eyepiece is a simple magnifier or, as it is colloquially called, a magnifying glass. The second lens is called the *objective* lens.

To understand the optical fundamentals of these instruments, it is necessary first to analyze the eyepiece (magnifier). Figure 20.16 illustrates an object placed near, but outside the focal point of a *magnifier.*

Without the magnifier, the object cannot be brought any closer than approximately 25 cm from the eye and still be clearly focused by the retina. The angle subtended on the retina is small, so the object appears small. When the magnifier is placed between the object and the eye, the object can be brought closer to the eye. The additional refraction caused by the magnifier subtends a larger angle on the retina (Figure 20.16), so the object appears enlarged.

Typically magnifier focal lengths are short, with values between 2 and 4 cm. A few selected rays from the endpoints of the object are shown. After refraction by the magnifier, these rays form an image, which appears to the eye, to be larger than the object. The object is "magnified."

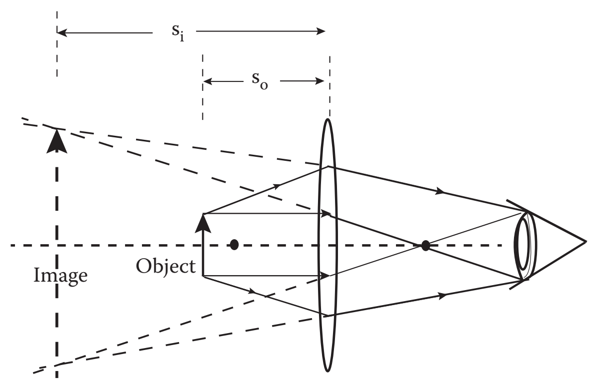


FIGURE 20.16 The image formed by a magnifier.

20.6.1 Refracting Telescope

The eyepiece (magnifier) is a convex lens of short focal length (f_e). The objective lens is also convex, but of a much larger focal length (f_0). These lenses are attached inside an otherwise light-tight tube. The rays entering the telescope from distant objects are nearly paraxial so the image formed by the objective lens is very near to its focal plane. The telescope is constructed so that this image, which is to be treated as the object of the magnifier, is located just inside the focal length of the eyepiece. The magnifier eyepiece forms an enlarged virtual image of this intermediate object. Thus, distant objects are enlarged. It can be shown that the magnification of this telescope, which is often called an astronomical telescope, is

$$M = -\frac{f_0}{f_e}. \tag{20.21}$$

20.6.2 Compound Microscope

The compound microscope, in contrast to the simple microscope, the magnifier, is so named because it consists of more than one lens. The fundamental microscope consists of two convex lenses, both of relatively short focal length. The object to be viewed is generally illuminated and placed just outside the focal length of the objective lens. Its image is real and inverted and is treated as the object of the eyepiece. The microscope is constructed so that the first image is located just outside the focal plane of the eyepiece. The eyepiece forms an image approximately 25 cm from it. Thus, the eye can focus clearly on this final image that is much larger than, and inverted, relative to the original object. It can be shown that the total magnification is

$$M = -\frac{25d}{f_0 f_e}. \tag{20.22}$$

Here, d is the separation distance between the eyepiece and objective lenses and is approximately equal to the length of the microscope.

PROBLEMS

20.1 The index of refraction of benzene is 1.50. Determine the speed of light in benzene.

20.2 An object is placed 12.0 cm on the reflecting side of a concave mirror. An image is formed on the same side at 30.0 cm from the vertex of the mirror. Calculate
 a. The focal length of the mirror.
 b. The height of the image if the object is 4.50 cm in height.

20.3 A light ray, in air, strikes the surface of a flat plate of glass at an angle of 55.0° to the normal. The refracted ray travels at 33.1° from the normal, inside the glass. Calculate the index of refraction of the glass.

20.4 The index of refraction of sodium chloride is n = 1.544. Determine the critical angle (θ_c) for total internal reflection in sodium chloride.

20.5 The focal length of a Plexiglas (n = 1.47) plano-convex lens, in air, is 10.0 cm. Calculate the radius of curvature (R_1) of the spherical side of the lens. (*Note:* A plano-convex thin lens has one flat (planar) side with (R_2 = infinity) and one spherically curved side with radius of curvature (R_1).)

20.6 An object is placed 9.00 cm to the left of a thin lens. The image is formed on the same side of the lens at 29.3 cm from it. Calculate
 a. The focal length of the lens.
 b. The height of the image if the object is 3.50 cm in height.
 c. Sketch a ray diagram of the refractions that locate the image position.

20.7 (a) A convex lens (f_1 = +12.0 cm) is situated to the left of a concave lens (f_2 = −12.0 cm). They are 46.0 cm apart. An object is placed 18.0 cm to the left of the convex lens. (a) Calculate the position of the final image for this lens combination. (b) Suppose now the lens separation is changed to 26.0 cm. All other parameters are unchanged. Determine the position of the final image for this altered lens combination.

20.8 Two rays of light originate and diverge from the same source. They are then reflected from a common plane mirror. The angle between the rays is 16.0° before reflection. Determine the angle between them after reflection.

20.9 (a) What is the shortest minimum length, of a plane mirror, in which a standing 6.00-ft man can see his total image? (b) How close to the mirror must he be?

20.10 A convex mirror has a radius of curvature R = 90.0 cm. An object 8.00 cm high is placed 24.0 cm in front of the mirror. Determine (a) the location and (b) size of the image formed.

20.11 A concave mirror has a focal length of 45.0 cm. An object, 8.00 cm high, is placed 24.0 cm in front of the mirror. Determine (a) the location and (b) the size of the image formed.

20.12 A glass window pane in a home has an index of refraction, n = 1.52, and is 5.00 mm thick. Determine (a) the velocity of light in the glass and (b) the time it takes for light to travel perpendicularly through the pane.

20.13 A 6.00-cm high object is placed 24.0 cm to the left of a convex lens (f = 16.0 cm). A concave mirror (f = 16.0 cm) is placed 60.0 cm to the right of the lens. Determine (a) the location (relative to the mirror), (b) the size, and (c) the nature of the final image.

20.14 A fish appears to be 1.10 m below the surface of a lake when viewed directly from above. Determine the fish's actual depth. *Note:* The index of refraction of water is n = 1.333.

20.15 Two diverging lenses are placed 32.0 cm apart. Each has a focal length of –16.0 cm. An object is placed 8.00 cm to the left of the leftmost lens. Determine the location of the final image.

20.16 The lenses in a compound microscope are 20.0 cm apart. The focal length of the objective lens is 1.30 cm. What value of focal length should the eyepiece be for the microscope to have a magnification of −85.0?

20.17 A refracting telescope has a magnification of −165. Its eyepiece has a focal length of 4.85 mm. Determine the focal length of the objective lens.

20.18 A lens has an index of refraction of $n_\ell = 1.52$ and a focal length of 18.0 cm when used in air. What would be the value of its focal length if it were completely immersed in water? *Note:* $n_w = 1.333$.

21 Physical (Wave) Optics

21.1 INTRODUCTION

Recall that geometrical optics deals with the properties of light when it encounters objects, mirrors, lenses, and orifices that are large compared with its wavelength. Physical or wave optics addresses light behavior when it encounters objects that are approximately the same size or that are smaller than its wavelength. It is in this realm that light exhibits its wave nature and a wave description of light is appropriate. Under these conditions, the two main manifestations of light behavior are called *interference* and *diffraction*. These two phenomena are consequences of a very general principle called the *linear superposition principle* (LSP).

21.1.1 Linear Superposition Principle

A wave is a periodic disturbance in a medium. The medium may be material, as in the case of sound waves traveling through air or a steel rod. Or, it may be nonmaterial as with electromagnetic waves (EMWs), where the medium consists of electric and magnetic fields. It is possible for two or more disturbances (waves) to meet at the same spatial point in a medium, at exactly the same time. Then, the question that arises is how will the medium respond, that is, what is the resultant disturbance at that spatial point and that particular time? The LSP states:

> The displacement (disturbance) at any point in the medium is equal to the algebraic sum of the individual displacements (disturbances).

For example, consider two EMWs whose oscillations are written as

$$E_1 = E_{01}\sin(k_1x_1 - \omega_1t_1) \tag{21.1}$$

and

$$E_2 = E_{02}\ \sin(k_2x_2 - \omega_2t_2). \tag{21.2}$$

If these two waves meet at some common point, that is, $x_1 = x_2 = P$, the resultant electric field at P, according to the LSP, is

$$E_{TOTAL} = \sum_i E_i = E_1 + E_2.$$

Figure 21.1a depicts two light rays (only the electric field components are shown for clarity) of equal wavelengths, emitted from two separate sources, propagating through air. Consider the point P that is located one and one-fourth (1&1/4) wavelengths from source one and two and one-fourth (2&1/4) wavelengths from source two. From Figure 21.1a, it is seen that waves one and two reach a maximum amplitude simultaneously. Similarly, they reach their zero and minimum amplitudes simultaneously, that is, they are "in step" with each other. They are said to be "in phase." Thus, the medium, which is an electric field, oscillates in step with the two disturbances and has an amplitude that is the arithmetic sum of the amplitudes of the two waves. In Figure 21.1a, the waves are

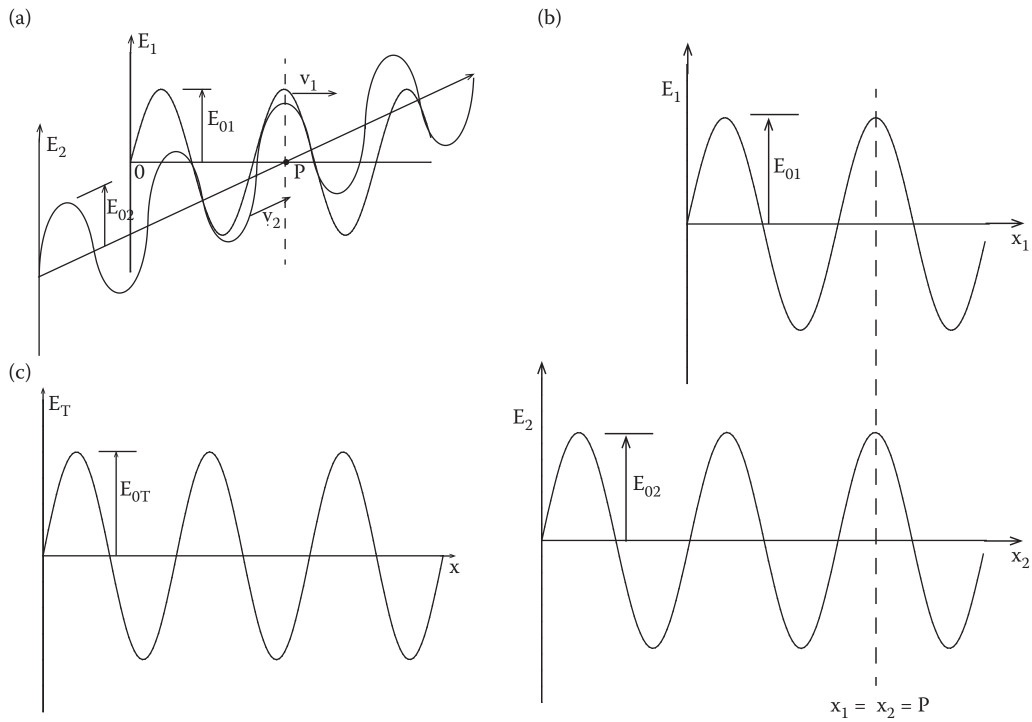


FIGURE 21.1 (a) Light rays emitted from two separate sources, (b) their relative phases, and (c) the superposition of the waves.

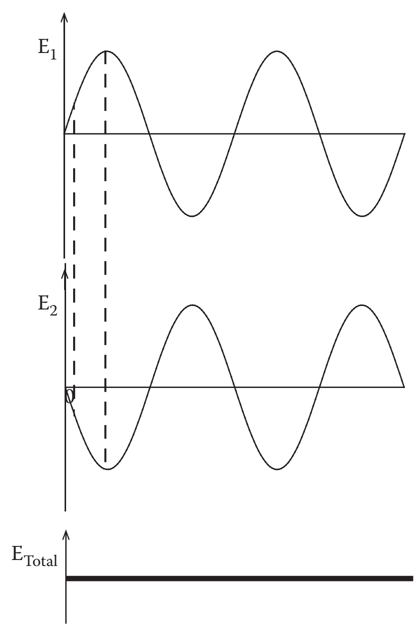


FIGURE 21.2 The relative phases for the total destructive interference of two waves.

traveling in separate directions for clarity. If they were traveling in the same direction, they might be depicted as in Figure 21.1b. The sum of their amplitudes is shown in Figure 21.1c. Since the wave disturbances interact spatially and temporally, they are said to "interfere" with each other, so this phenomenon is known as *interference*.

The special case, depicted in Figure 21.1c where the waves are perfectly in phase, is called *total constructive interference*. Correspondingly, if for example, source two in Figure 21.1a was moved one-half (1/2) wavelength closer to point P, its wave would be out of step with wave one. Indeed, wave two would have a minimum amplitude at the time wave one had a maximum amplitude. In colloquial jargon, "one's pushing while the other's pulling." This is called *total destructive interference* and is illustrated in Figure 21.2.

It is clear that point-by-point signals E_1 and E_2 add to $E_{TOTAL} = 0$. If, at any point P, the waves are not totally in phase, that is where they both reach maximum amplitude at the same time, or are totally out of phase, where one is at a maximum amplitude when the other is at its minimum amplitude, then interference still occurs, but it is neither totally constructive nor totally destructive.

21.2 DOUBLE-SLIT EXPERIMENT

The earliest (known) demonstration of interference and the wave nature of light was the double-slit experiment performed in 1801 by a brilliant English scientist Thomas Young. To demonstrate interference, two or more *coherent sources* of light are required. If the light waves emitted from the sources maintain a constant phase relation relative to each other, the sources are coherent. The white light emitted from most sources is not coherent because of the many different wavelengths, which constitute it. Young devised an ingenious way to acquire two coherent sources from white light. Today, we have lasers that emit monochromatic and therefore coherent light waves.

Consider the double-slit apparatus, as viewed in Figure 21.3. S is a laser that emits coherent light. Recall that wavefronts, by definition, represent the locus of points of like phase. The emitted wavefronts impinge on an opaque screen with two very narrow vertical slits labeled s_1 and s_2 that are located equidistant from the laser.

By symmetry, the two points on a *given* wavefront that reaches the slits will be in phase. If light possesses a wave nature, the wavefronts at s_1 and s_2 will simultaneously spread out after passing through the slits. Thus to the opaque screen, located a distance L from the slit-screen, s_1 and s_2 behave as two coherent sources of light. The light from the two sources will spread out, to the right of the slit-screen, and, by the LSP, interfere with each other, to produce bright and dark regions (*fringes*) on the screen. Where, on the screen, will the fringes appear? The brightest fringes will

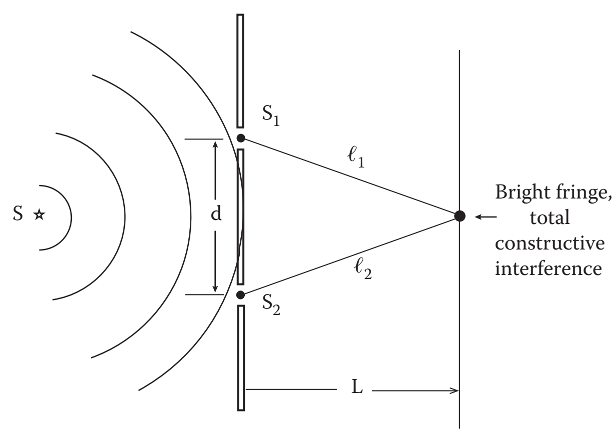


FIGURE 21.3 The Young's double-slit apparatus.

occur for total constructive interference, the darkest for total destructive interference. The regions between these two extremes will be neither bright nor totally dark.

Consider the bright fringes. If the difference in distances, ℓ_1 and ℓ_2, traveled by the light from s_1 and s_2 is simultaneously equal to an integer number of wavelengths of the light, total constructive interference occurs. In Figure 21.3, $\ell_1 = \ell_2$ so $\Delta\ell = \ell_1 - \ell_2 = 0$, thus there is constructive interference and a bright central fringe exists.

Figure 21.4 shows an improved apparatus where the laser source is placed at the focal point F of a converging lens. After refraction by the lens, the rays are paraxial and the wavefronts, by definition, are perpendicular to the rays.

Again, the light passing through the slits behaves as two coherent sources. At various locations on the screen, the light from s_1 will have traveled a different distance ℓ_1, than the light from s_2. Constructive interference, and hence a bright fringe, will occur whenever the path difference $\Delta\ell$ equals an integer number of wavelengths, that is,

$$\Delta\ell = m\lambda, \quad m = 0, 1, 2, 3, \ldots \tag{21.3}$$

The integer m is called the *order* of the bright fringes. For example, m = 1 is a first-order fringe, m = 2 a second-order fringe, and so on.

Let θ_B (Figure 21.4) represent the angles at which bright fringes occur. If the triangle (s_1 ps_2) is approximated by a right triangle, then $\Delta\ell = d \sin \theta_B$ and Equation 21.3 can be written as

$$d \sin \theta_B = m\lambda, \quad m = 0, 1, 2, 3, \ldots \tag{21.4}$$

For dark fringes at $\theta_{D,}$ the path difference must be an odd, half-integer of the wavelength, that is,

$$\Delta\ell = \left(m + \frac{1}{2}\right)\lambda, \quad m = 0, 1, 2, 3, \ldots \tag{21.5}$$

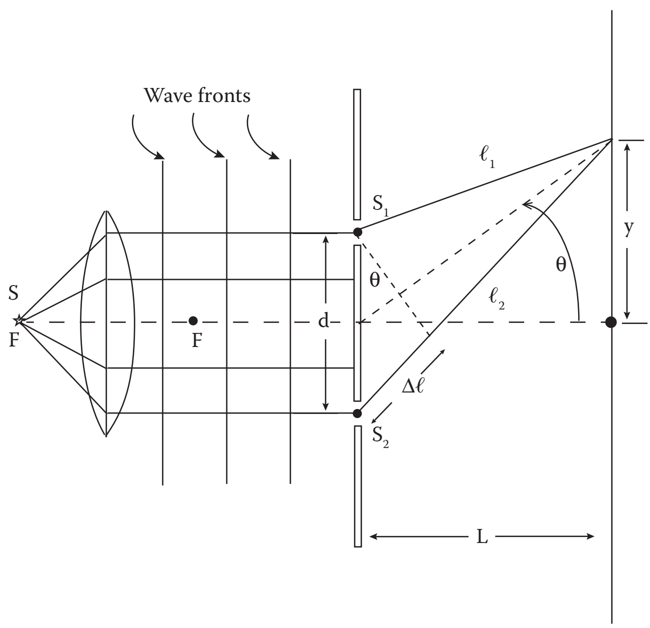


FIGURE 21.4 An improved double-slit apparatus.

or

$$d \sin \theta_D = \left(m + \frac{1}{2}\right)\lambda. \tag{21.6}$$

The location of the bright or dark fringes (y), relative to the central bright fringe, can also be found. In Figure 21.4, it is seen that

$$y = L \tan \theta.$$

If $L \gg y$ then $(y/L) \ll 1$ and the $\tan \theta \ll 1$ can be approximated by $\sin \theta$. That is,

$$\tan \theta \approx \sin \theta. \tag{21.7}$$

For bright fringes, using Equations 21.6, 21.7, and 21.4,

$$y_B \cong L \sin \theta_B = m\left(\frac{L\lambda}{d}\right). \tag{21.8}$$

Similarly, for dark fringes

$$y_D \cong \left(m + \frac{1}{2}\right)\frac{L\lambda}{d}. \tag{21.9}$$

EXAMPLE 21.1

Two optical slits (double-slit apparatus) are separated 2.00×10^{-4} m apart and illuminated by orange light, $\lambda = 625$ nm. A screen is placed 1.75 m from the slits. Determine (a) the angle of the second bright fringe and (b) the distance, along the screen, between the third and fist bright fringes.

Solution

a. Using Equation 21.4 with m = 2 gives

$$\sin \theta_B = m\left(\frac{\lambda}{d}\right) = 2\left(\frac{6.25 \times 10^{-7}\ \text{m}}{2.00 \times 10^{-4}\ \text{m}}\right) = 6.25 \times 10^{-3}$$

$$\theta_B = \sin^{-1}(0.00625) = 0.358^{\circ}.$$

b. Using Equation 21.8 with m = 3, then m = 1, and subtracting to find the difference between the fringes, gives

$$y_3 - y_1 = (3-1)\left(\frac{L\lambda}{d}\right) = (2)\left(\frac{(1.75\ \text{m})(6.25 \times 10^{-7}\ \text{m})}{2.00 \times 10^{-4}\ \text{m}}\right) = 0.0109\ \text{m} = 1.09\ \text{cm}.$$

EXAMPLE 21.2

Two narrow optical slits are 2.00×10^{-4} m apart and situated 1.50 m in front of an opaque screen. The slits are illuminated with coherent light and the center of the second bright fringe is located 8.85 mm from the center of the central maximum fringe. Determine the wavelength of the light.

Solution

Equation 21.8 gives the distance of the mth-order bright fringe from the central maximum. So,

$$y_B = m\left(\frac{\lambda L}{d}\right)$$

or

$$\lambda = \frac{y_B d}{mL} = \frac{(8.85 \times 10^{-3}\ \text{m})(2.00 \times 10^{-4}\ \text{m})}{(2)(1.50\ \text{m})} = 5.90 \times 10^{-7}\ \text{m} = 590\ \text{nm}$$

21.3 THIN-FILM INTERFERENCE

The pattern of colors that are seen reflected from a thin film of oil or gasoline, floating on the water in your driveway on a rainy day, is an example of thin-film interference. A portion of the ambient white light from the sun incident on the film is reflected from the top surface of the oil and some part is transmitted through the film and reflected from the bottom surface. These two rays, considered separate waves, obey the LSP and interfere with each other on their travel to your eye. The type of interference, constructive or destructive, will depend on the relative phases of the two waves, which in turn depend on their path difference, wavelength, and phase changes upon reflection. Phase changes at reflection requires some additional discussion.

Recall that if a traveling mechanical wave moves from one medium to another, at the interface between the two media, part of the wave will be transmitted and part reflected. The phase of the transmitted wave will be unaltered, but the phase of the reflected wave may be changed. If the wave travels from a higher mass density medium to one of lower mass density, the reflected wave will not undergo a phase change. But, if it travels from lower to higher density, the reflected wave will undergo a 180° phase change at the interface.

The same phenomenon occurs with light waves, but it is the optical density, that is, index of refraction, as opposed to the mass density, that determines the phase change upon reflection. Figure 21.5 depicts a single pulse of a light wave going from a medium of index of refraction n_1 into a medium of index n_2.

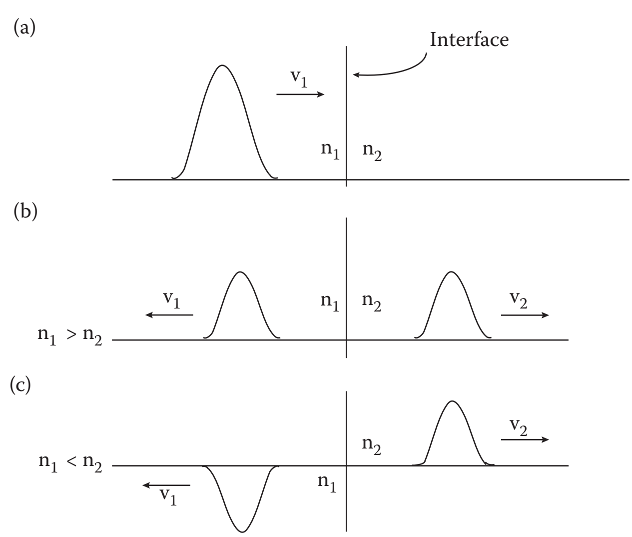


FIGURE 21.5 (a) A pulse incident on an optical interface, (b) a reflected pulse for $n_1 > n_2$, and (c) a phase-inverted reflected pulse when $n_1 < n_2$.

In Figure 21.5a, the pulse is approaching the interface. Figure 21.5b shows the transmitted and reflected pulses a short time after the incident pulse encounters the interface for the case $n_1 > n_2$. Figure 21.5c is for the case $n_1 < n_2$, where the reflected pulse undergoes a 180° phase shift. Note that transmitted pulses are not phase shifted.

A single pulse and its shape, a short time after encountering the interface, are shown here for clarity. Of course, a light wave consists of continuous oscillations, that is, a continuum of pulses with both positive and negative segments, so the phase shifts shown in Figure 21.5 are occurring continuously.

Now consider a thin film of thickness d, surrounded by air and illuminated by white light at approximately normal incidence (Figure 21.6). If the film is a thin piece of, say, crown glass ($n_2 = 1.532$), then $n_2 > n_1 = n_{air}$.

The part of the incident ray reflected from the top surface will be phase shifted by 180° since $n_1 < n_2$. Part will be transmitted through the top surface, with no phase shift, and be reflected from the bottom surface. This ray will not be phase shifted on reflection but will travel upward through the film where part of it will be transmitted through the upper surface, again unshifted.

For constructive interference to occur, for a given wavelength, rays 1 and 2 must be in phase when they enter the eye and consequently at the top surface of the film. For this to happen, an odd integer number of half wavelength must fit into the film path (abc) ~2d. The additional $(1/2)\lambda$ is to account for the phase shift of ray 1. Thus,

$$2d = \left(m + \frac{1}{2}\right)\lambda_{\text{medium}}, \quad m = 0, 1, 2, \ldots. \tag{21.10}$$

Now,

$$n_{\text{med.}} \equiv \frac{c}{v_{\text{med.}}} = \frac{\lambda_{\text{air}} f_{\text{air}}}{\lambda_{\text{med.}} f_{\text{med.}}} = \frac{\lambda_{\text{air}}}{\lambda_{\text{med.}}}. \tag{21.11}$$

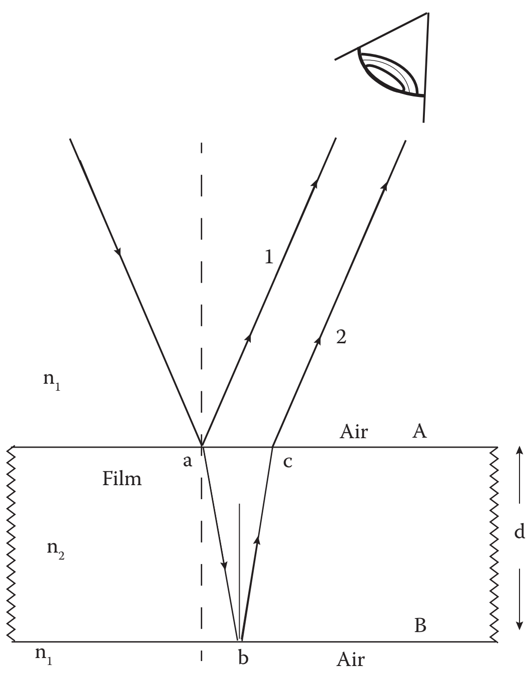


FIGURE 21.6 Interference of light by a thin transparent film.

The frequency is just the "counting" of the oscillations per second which does not change as the wave travels from one medium to the next. Using Equation 21.11, Equation 21.10 can be rewritten as

$$2nd = \left(m + \frac{1}{2}\right)\lambda_{air}, \quad m = 0, 1, 2, \ldots \tag{21.12}$$

For total destructive interference, the rays reflected from the top and bottom surfaces of the film must be 180° out of phase at the top surface and must be approximately of equal amplitude. If the film is only several wavelengths thick, the ray reflected from the bottom (ray 2) will not have lost appreciable energy in traversing the film twice. Thus, its amplitude will be approximately equal to ray 1's amplitude. Under these assumptions,

$$2d = m\lambda_{med.} = m\left(\frac{\lambda_{air}}{n}\right) \tag{21.13}$$

or

$$2nd = (m)\lambda_{air}, \quad m = 0, 1, 2, \ldots \tag{21.14}$$

Equation 21.14 gives the condition for destructive interference.

EXAMPLE 21.3

A thin film of gasoline floats on a small puddle of water in a driveway. Illuminated by white light, from the sun overhead, the film has a red appearance to an observer looking down on it. Determine the minimum nonzero thickness of the film. *Note:* $\lambda_{red} = 660$ nm, $n_{gas} = 1.40$, and $n_{water} = 1.33$.

Solution

The red appearance means constructive interference has occurred for red, thus enhancing it, while the other colors of the white light spectrum are not enhanced. Note that there are three media involved: air, gasoline, and water. The gasoline, with $n_{gas} = 1.40$, floats on water, $n_{water} = 1.33$, so there is no phase shift of red light reflected from the gasoline water (bottom) interface. Thus, Equation 21.12 can be used directly. For $m = 1$,

$$2n_{gas}\, d = \left(1 + \frac{1}{2}\right)\lambda_{air},$$

so,

$$d = \frac{(3/2)\lambda_{air}}{2n_{gas}} = \frac{(3/2)(660\ \text{nm})}{2(1.40)} = 354\ \text{nm} = 3.54 \times 10^{-7}\ \text{m}.$$

EXAMPLE 21.4

An antireflection material ($n_{coat} = 1.35$) is to be coated onto a flint glass ($n_{lens} = 1.66$) binocular lens to enhance transmission, that is, minimize reflection of green light ($\lambda = 550$ nm) by the lens. Determine the minimum thickness of the coating.

Solution

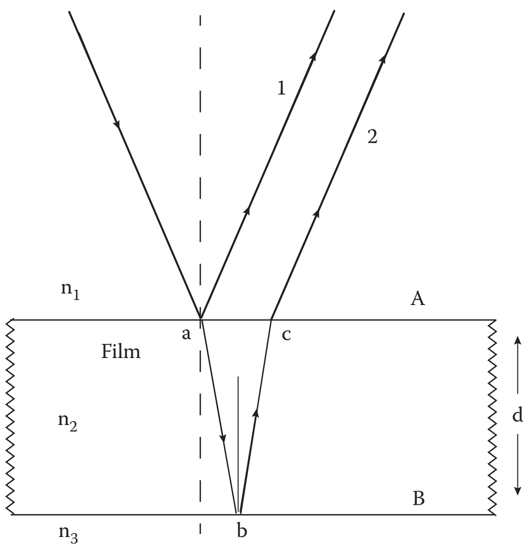
The coating should cause total destructive interference, that is, no partial cancelation of the incident green light by the reflected green light. Since $n_{air} < n_{coat}$, the light will undergo a 180° phase shift at the air-coating interface. Also, $n_{coat} < n_{lens}$, so a 180° phase shift will also occur at the coating–lens interface. Thus, phase shifts occur at *both* interface reflections. For destructive interference, the light reflected from the second interface must be phase shifted by (1/2)λ, so applying Equation 21.12

$$2n_{coat}\, d = \left(m + \frac{1}{2}\right)\lambda_{air}$$

or

$$d = \frac{(3/2)\lambda_{air}}{2n_{coat}} = \frac{(3/2)(550\text{ nm})}{2(1.35)} = 306\text{ nm}.$$

A word of caution is appropriate regarding the use of Equations 21.12 and 21.14. The figure below shows the thin film of index n_2, surrounded by media of indices n_1 and n_3 and the two interfaces A and B.



Thin-film interference when the film is backed by a third, light transparent medium.

If a phase shift occurs at only one interface, A or B, Equation 21.12 is the condition for constructive interference and Equation 21.14 for destructive interference. Alternately, if a phase shift occurs at both ($n_1 < n_2 < n_3$) or neither ($n_1 > n_2 > n_3$) surfaces, Equation 21.12 is the condition for destructive, not constructive, interference and Equation 21.14 for constructive interference. So, analysis of phase shifts is necessary to determine which of Equations 21.12 or (21.14) is appropriate to use. Recall that whether a shift occurs or not depends on the relative values of the indices of refraction on both sides of the interface.

If a thin film is not of uniform thickness, possibly wedge shaped in cross section, interference can still occur. The conditions for constructive or destructive interference, Equations 21.12 and 21.14, depend on both film thickness d and wavelength λ_{air}. The film will enhance a given color (λ_{air}) in those regions of the film where d satisfies Equation 21.12. We have all observed these different colors in various regions of a nonuniform film.

21.4 SINGLE-SLIT DIFFRACTION

Diffraction is the bending of light waves as they pass by obstacles whose size is comparable to their wavelength. The obstacles may be opaque (nontransmitting) or transparent (transmitting) to the waves. Diffraction of *sound* waves is more easily observed than diffraction of *light* waves. As an example of sound diffraction, if you are in a room connected to a second room by an open doorway and a person, who you cannot see, is in the second room talking, you can hear them. The sound waves they create, say at 500 Hz, have a wavelength of approximately 0.68 m = 68 cm = 27 inches (v_{sound} = 340 m/s). The width of a typical interior doorway is 32 inches. The "obstacles," that is, "slit" created by the doorway and the sound wavelength, are of comparable size, so the "bending" is easily observable.

For diffraction of light, consider a single narrow slit of width a, cut in an opaque plane, and viewed as in Figure 21.7.

The source of the light S is placed at the focal point F of a converging lens. After refraction, the rays are paraxial. The wavefronts are perpendicular to the rays and all points on a given wavefront are in phase. Note that the time of travel from source through the lens, for the three rays shown, is equal; that is, the three *optical* paths are equal. Thus, the lens does not cause a phase difference between the rays.

The wavefronts pass through the slit and illuminate a screen located a distance L from the slit plane. At any given time, any point P on the screen is illuminated by light coming from all points on the wavefronts that pass through the slit. Since various small segments of the wavefronts are at different distances from P, the light coming from these segments will interfere with each other. So diffraction is really an interference phenomenon, that is, interference of light coming from different regions on the same wavefront. The interference creates bright and dark regions (fringes) on the screen that is called correspondingly, a *diffraction pattern*.

It is easier to analyze the dark regions since they occur when there is total destructive interference from all segments of the wavefront. Semibright fringes (called secondary maxima) occur when there is more constructive, but not total, interference than destructive interference. They are not discussed here.

The condition for the first dark fringe can be acquired by envisioning the slit, of width (a), as two halves, each of width (a/2). Also, consider each point on the wavefront as a source of light that propagates outward toward the screen. The first dark fringe occurs when light from the top of the top half (point A) interferes destructively with light from the top of the bottom half (point B). If

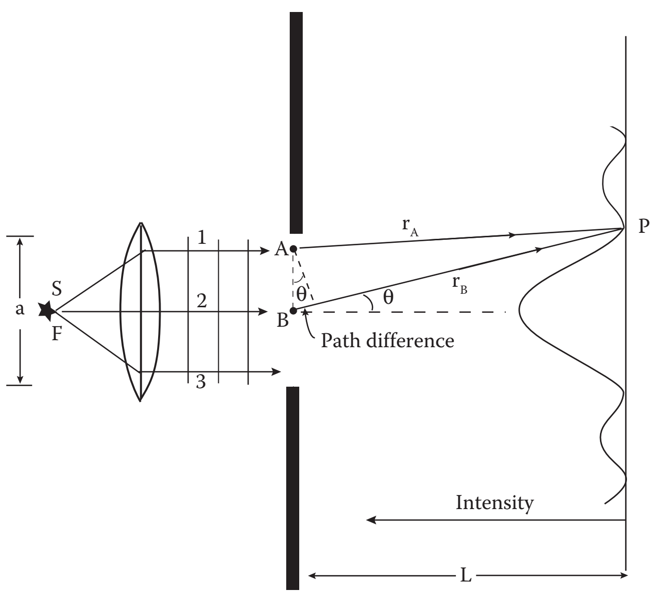


FIGURE 21.7 The apparatus for single-slit diffraction.

the path difference, to the screen, of these two sources is (1/2)λ, they will destructively interfere. Continue on down the two halves of the wavefront, in the slit, such that *pairs* of points, one from the top half with a corresponding one from the bottom half, interfere destructively. Thus,

$$\text{Path difference} = (r_B - r_A) = \left(\frac{1}{2}\right)\lambda. \tag{21.15}$$

From Figure 21.7,

$$\text{Path difference} = \left(\frac{a}{2}\right)\sin\theta_D, \tag{21.16}$$

where θ_D is the angle at which the dark fringe appears. Combining Equations 21.15 and 21.16 gives

$$\left(\frac{1}{2}\right)\lambda = \left(\frac{a}{2}\right)\sin\theta_D$$

or

$$\lambda = a\sin\theta_D. \tag{21.17}$$

To determine the location of the second dark fringe, subdivide the slit into four equal segments of length (a/4) such that there is cancelation (destructive interference), by pairs of points on the segmented wavefront.

So,

$$\text{Path difference} = \left(\frac{1}{2}\right)\lambda = \left(\frac{a}{4}\right)\sin\theta_D$$

or

$$2\lambda = a\sin\theta_D. \tag{21.18}$$

The third dark fringe is determined by subdividing the slit into six equal segments, each of length (a/6), and so on.

Generalizing, the condition for dark fringes is

$$m\lambda = a\sin\theta_D, \quad m = 1, 2, 3, \ldots. \tag{21.19}$$

In Figure 21.7, it can be seen that a maximum light intensity (energy/area/time) occurs at the point O on the screen. This central point is equidistant from points on the wavefront that are paired by starting at the top of the slit and at the bottom of the slit and simultaneously pairing toward the slit's center. The light from these paired points constructively interfere with each other to give the maximum fringe at point O. The relative light intensity on the screen is also plotted in Figure 21.7.

EXAMPLE 21.5

A narrow slit of width $a = 3.00 \times 10^{-6}$ m is illuminated by monochromatic light of $\lambda = 500$ nm. (a) Calculate the angle of the first-order dark fringe (m = 1), (b) its location on the screen, if L = 1.20 m, and (c) the wavelength of light needed, using the same slit, to place the first dark fringe at $\theta_D = 13.0°$.

Solution

a. Using Equation 21.19

$$m\lambda = a \sin \theta_D$$

or

$$\sin \theta_D = \frac{m\lambda}{a} = \frac{(1)(5.00 \times 10^{-7}\ \text{m})}{(3.00 \times 10^{-6}\ \text{m})} = 0.167$$

$$\theta_D = \sin^{-1}(0.167) = 9.61°.$$

b. From Equation 21.8

$$\tan \theta_D = \left(\frac{y}{L}\right)$$

$$y = L \tan \theta_D = (1.20\ \text{m})\tan(9.61°) = (1.20\ \text{m})(0.169) = 0.203\ \text{m} = 20.3\ \text{cm}.$$

c.

$$\lambda = \left(\frac{a}{m}\right)\sin \theta_D = \left[\frac{3.00 \times 10^{-6}\ \text{m}}{1}\right]\sin(13.0°) = 6.75 \times 10^{-7}\ \text{m} = 675\ \text{nm},$$

this is orange-red.

One can appreciate what Equation 21.19 implies by inserting special case values into it. Suppose $m = 1$, then

$$\sin \theta_D = \left(\frac{\lambda}{a}\right).$$

If $\lambda = a$, which would be a very narrow slit, then $\sin \theta_D = 1.0$ and $\theta_D = 90°$. This result gives the location of the first dark fringe at 90° which implies that the whole screen is illuminated, that is, light fills the space after passing through the slit.

If, on the other hand, $a = 100\lambda$, then

$$\theta_D = \sin^{-1}\left(\frac{1}{100}\right) = 0.573°.$$

The second dark fringe would appear at

$$\theta_D = \sin^{-1}\left(\frac{m\lambda}{a}\right) = \sin^{-1}\left[\frac{(2)\lambda}{(100)\lambda}\right] = \sin^{-1}\left(\frac{1}{50}\right) = 1.15°.$$

The dark fringes would appear very close to the central maximum, so the bending of the light would be hard to observe. This is why sound diffraction is much more apparent than light diffraction.

If, instead of a long narrow slit, the single aperture was a circle, for example, a pinhole in an opaque object, the diffraction pattern would appear as circular fringes.

21.5 DIFFRACTION GRATING

Patterns of light and dark fringes are seen when plane monochromatic light passes through a single slit or a double slit. Such patterns are also observed when light passes through many slits.

Consider a planar opaque material with many closely but equally spaced slits (Figure 21.8). This arrangement of slits is called a *plane diffraction grating*. Only five slits are shown in Figure 21.8, but industrial and research grade gratings may have as many as 40,000 slits/cm. In Figure 21.8, the quantity d represents the distance between slits, that is, the slit separation. With a large number of slits per centimeter, the slits will be narrow enough so that the diffracted light from a single slit spreads out enough to interfere with the diffracted light from all the other slits. The light from the slits illuminates the screen (Figure 21.8). A converging lens is placed between the grating and screen so that its focal plane coincides with the screen. For rays that meet at a point P on the screen, the lens ensures that they were parallel after passing through their respective slits. The path lengths of these parallel rays that meet at point P are unequal. Recall that the optical path lengths of these parallel rays, passing through the lens, are equal. Any difference in path lengths occurs before the rays enter the lens. From Figure 21.8, the path difference between rays passing through adjacent slits is

$$\Delta s = d \sin \theta.$$

For constructive interference, this path difference must be equal to an integer number of wavelengths, that is,

$$m\lambda = d \sin \theta_B, \quad m = 0, 1, 2, 3, \ldots. \tag{21.20}$$

For destructive interference, the path difference Δs must be an odd half-integer number of wavelengths or

$$\left[m + \left(\frac{1}{2}\right)\right]\lambda = d \sin \theta_D, \quad m = 0, 1, 2, 3, \ldots. \tag{21.21}$$

From Equation 21.20, it is clear that the angular location of the bright fringes θ_B depends on the wavelength of the light incident on the grating. If white light, for example, from the sun, which contains all visible wavelengths, falls on the grating, a rainbow-like spectrum of colors would be

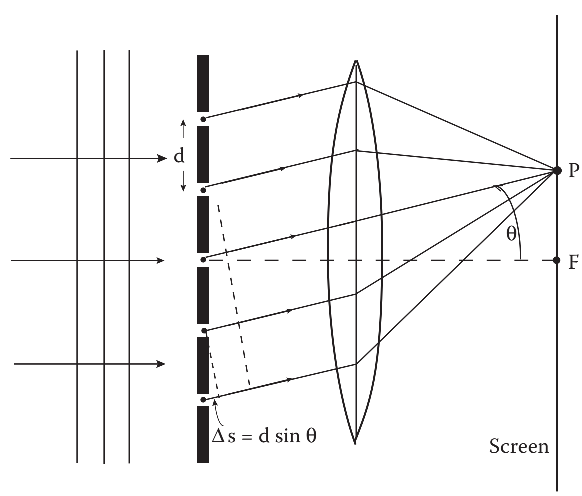


FIGURE 21.8 Interference by a plane “diffraction” grating.

observed on either side of the central maximum. The central maximum (m = 0) would be white because all colors constructively interfere at $\theta_B = 0$.

EXAMPLE 21.6

White light is incident on a diffraction grating that has 1.02×10^4 lines/cm. Determine the angles (θ_B) that locate the first-order maxima of the endpoint wavelengths $\lambda_{violet} = 690$ nm and $\lambda_{red} = 410$ nm.

Solution

Equation 21.20 can be used once the slit separation d is determined.

$$d = \frac{1}{1.02 \times 10^4 \text{ lines/cm}} = 9.80 \times 10^{-5} \text{ cm/line},$$

$$m\lambda = d \sin \theta B,$$

so,

$$\theta_{violet} = \sin^{-1}\left[\frac{m\lambda}{d}\right] = \sin^{-1}\left[\frac{(1)(6.90 \times 10^{-7} \text{ m})}{9.80 \times 10^{-7} \text{ m}}\right] = \sin^{-1}(0.704) = 44.7°$$

and

$$\theta_{red} = \sin^{-1}\left[\frac{m\lambda}{a}\right] = \sin^{-1}\left[\frac{(1)(4.10 \times 10^{-7} \text{ m})}{9.80 \times 10^{-7} \text{ m}}\right] = \sin^{-1}(0.418) = 24.7°.$$

A useful instrument called a *grating spectrometer* can be constructed using a diffraction grating, a converging lens to collimate incoming light, and a telescope to view the grating fringes (see Figure 21.9).

These items are mounted on a fixed table with the telescope free to rotate about the grating, which is located at the table center. The spectrometer can be used to help identify unknown materials. As an example, the light from a gas(vapor) that is incandescent consists of discrete wavelengths that are unseparated as they enter the unaided eye. If this light, instead, enters the slit of the spectrometer, it is made paraxial by the collimating lens and is then incident on the grating. The grating "sorts out" the discrete wavelengths whose angles are measured with the telescope. By Equation 21.20, the

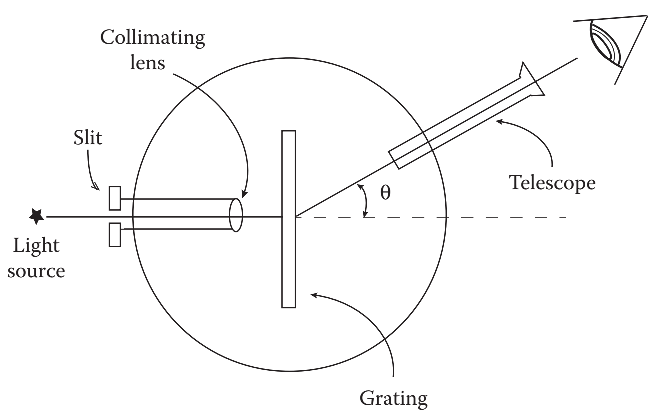


FIGURE 21.9 A grating spectrometer "sorts" light according to wavelength.

wavelengths can be determined. These wavelengths are characteristics of electron energy transitions in certain atoms, so their presence helps identify the presence of such atoms in the incandescent gas.

Another, more modern application of physical optics (interference) is found in the operation of compact disks (CD) and digital video disks (DVD). Information (sound and pictures) is stored on the disk in a combination of very small flat protrusions (called "pits") separated by flat regions of no protrusions (called "land"), both arranged along a spiral track. The pits and land are covered with a flat transparent coating of plastic. As the disk rotates, a stationary laser beam is reflected from the bottom surface and into a light detector. The reflected light varies in intensity as the pits and land cross the beam. These variations are sensed by a stationary detector and converted into fluctuating voltages that are transformed into binary numbers, that is, "zeros" and "ones." The variations in reflected light are maximized if the pit thickness d causes total destructive interference. How does this work? As a moving pit starts to be irradiated by the laser, part of the beam is reflected by the land and part by the pit. This happens at the edge of the pit. The part reflected from the land must travel farther by an additional distance 2d than the part reflected by the pit. The thickness of the pit is chosen so that 2d equals one-half the wavelength of the laser beam in the plastic coating. Thus, the part of the beam reflected from the land is 180° out of phase with the part reflected from the pit. At that instant, the two parts of the reflected beam exhibit destructive interference when they combine. The detector thus "sees" essentially zero intensity reflections at the edges of the pits and greater reflected intensity from the uniformly flat surfaces. These fluctuations are converted into digital signals.

21.6 POLARIZATION

Light waves, EMWs in general, are *transverse* waves. Recall that a wave is a periodic oscillation in a material or nonmaterial medium. For light waves, the oscillating media are electric and magnetic fields that oscillate *perpendicular* to the direction of wave propagation (Figure 19.4). For longitudinal waves, the oscillations are parallel to the wave velocity vector.

The word *polarization*, as used in this section, pertains to the spatial orientation of the electric and not to the magnetic field of an EMW. Recall that, spatially, the magnetic field oscillation is perpendicular to both the electric field and the propagation direction. Figure 19.4 shows the oscillation of **E** in a plane, the xy plane. Such light is called *plane polarized* or, sometimes, linearly polarized.

Light can be produced by accelerating charged particles, but most light originates from electron energy transitions in atoms. The pulse emitted by this process is polarized in a fashion similar to EMW emission from a simple radio antenna (Figure 19.3). A later pulse emitted by this atom may have a different polarization direction. Additionally, atoms in a typical light source (lasers excluded, here) act independently and emit various and random polarization directions. By the LSP, the sum of these pulses, by the collection of emitting atoms, constitutes the rays of light emitted by the source. These rays have a continuum of polarization directions, so the light is *unpolarized*. Figure 21.10

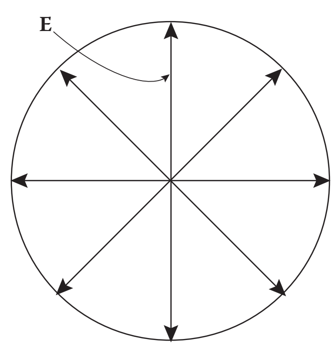


FIGURE 21.10 Unpolarized light consists of **E** vectors oscillating in all planes perpendicular to the wave propagation direction.

depicts some **E** vectors that constitute an unpolarized ray as viewed along its direction of propagation. Figure 21.11 is a side view of the same unpolarized ray.

Note: A light ray with the plane of **E** in some arbitrary position is shown in Figure 21.12a. The **E** vector may be resolved into components E_x and E_y as shown in Figure 21.12b. Thus, unpolarized light can be described in just these two components. The infinite number of planes of **E** about the propagation direction is not needed in the description.

There are ways, to be discussed shortly, to eliminate or remove most of the continuum of polarization directions so that only one oscillation direction, in a single plane, remains. Such light is plane polarized (see Figure 21.13).

Plane-polarized light can be produced from unpolarized light by certain materials. Such materials absorb all electric field components except those in a certain plane. The direction of **E** transmitted defines the *transmission axis* of the polarizing material, called the *polarizer* (Figure 21.14).

The polarizer can also be used as an *analyzer*, as in Figure 21.15. Note that in Figure 21.15, the **E** field plane of the incident, polarized light is not in the same direction as the transmission axis of the analyzer. So, only the amount (**E** cos θ) gets through the polarizer/analyzer. Since the intensity of EMWs is proportional to the square of the electric field, the average intensity leaving the analyzer is

$$(S)_{ave} = (S_0)_{ave} \cos^2(\theta). \tag{21.22}$$

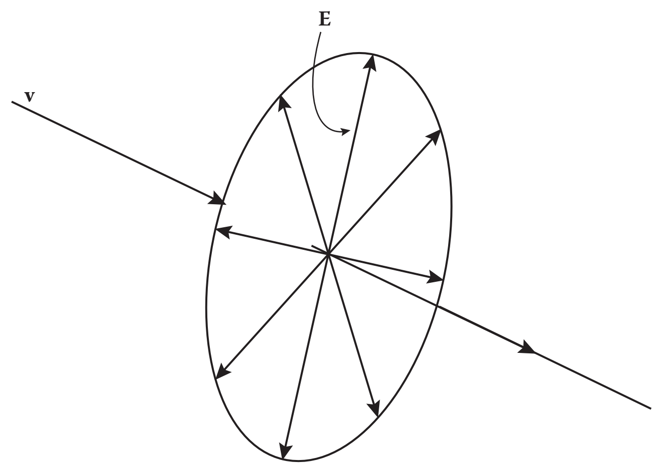


FIGURE 21.11 A side view of a plane of unpolarized light.

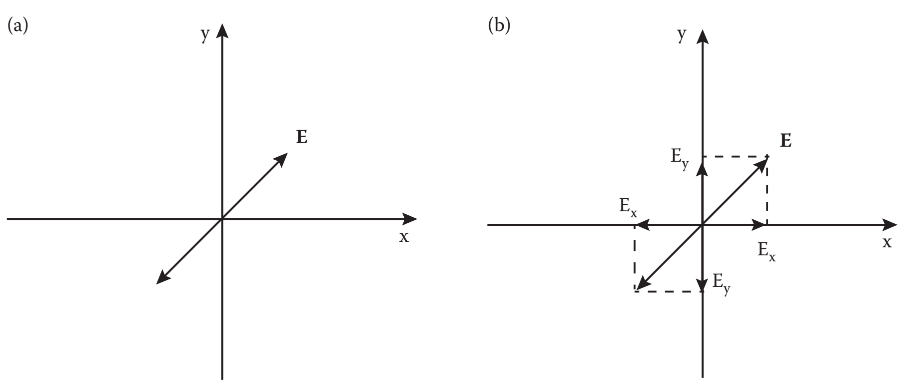


FIGURE 21.12 (a) The plane of **E** in some arbitrary direction and (b) the arbitrary **E** vector resolved into x and y components.

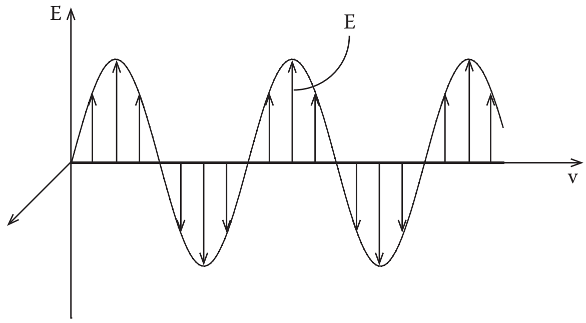


FIGURE 21.13 In plane-polarized light, **E** oscillates in a single plane.

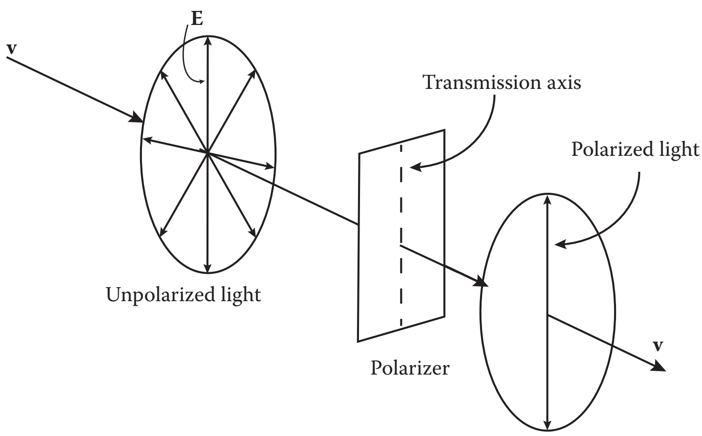


FIGURE 21.14 Plane-polarized light extracted from unpolarized light.

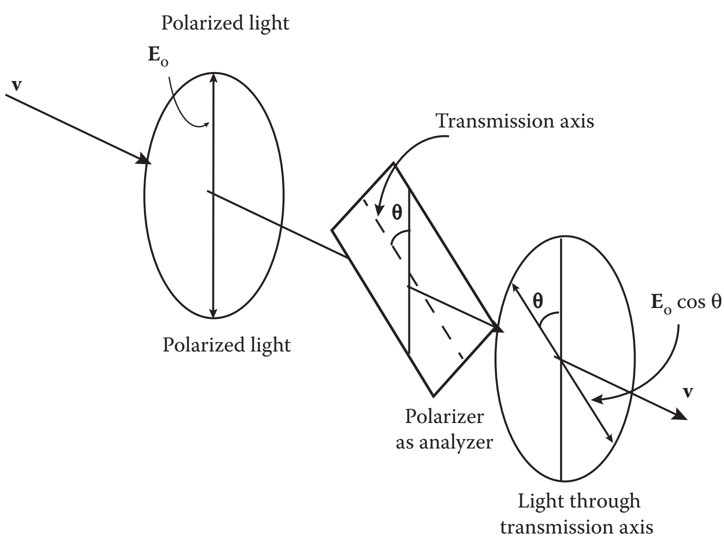


FIGURE 21.15 A polarizer can be used to analyze the plane of **E**.

Equation 21.22 is known as *Malus' law*, after its discoverer, the French engineer Louis Malus (1775–1812).

EXAMPLE 21.7

Suppose a photodetector is placed immediately after (downstream) the analyzer in Figure 21.15. At what angle θ should the transmission axis be oriented so that the intensity of the incident beam is reduced to one-third of its average value?

Solution

It is given that

$$(S)_{ave} = \frac{1}{3}(S_0)_{ave}.$$

Malus' law gives

$$(S)_{ave} = \frac{1}{3}(S_0)_{ave} = (S_0)_{ave}\cos^2\theta$$

or

$$\cos\theta = \left[\frac{1}{3}\right]^{1/2} = (0.333)^{1/2} = 0.577,$$

so,

$$\theta = \cos^{-1}(0.577) = 54.8°.$$

Polarized light can be produced by several processes. They are

1. Selective absorption
2. Reflection
3. Double refraction
4. Scattering

A brief discussion of each follows.

Absorption: The American inventor, Edwin Land (1909–1991) constructed a material (Trade named; *Polaroid*), in 1932, that consists of long straight chains of polymer molecules. These molecules are good electrical conductors along their long axes and very poor conductors across their thin widths. This property is called *dichroism*. He was able to align these long molecules, like a picket fence, by solidifying them between cellulose sheets. When unpolarized light passes through these translucent sheets, the components of **E** that are parallel/antiparallel to the long axes of the polymer molecules are absorbed by the electrical conduction process. The perpendicular components of **E** are not absorbed and pass through the Polaroid. Thus, the transmitted light is polarized.

Indeed microwaves, that occupy a different wavelength rang of the EMW spectrum, can be polarized by a similar process. If they pass through a planar array of closely spaced conducting wires, the transmitted microwaves are polarized (see Figure 21.16). Notice that the polarization direction is perpendicular to the long axes of the molecules/wires.

Reflection: Unpolarized light reflected from a surface may be partially or totally polarized. If the reflecting material is transparent or translucent so that part of the incident light is refracted, both reflected and refracted rays will be partially polarized.

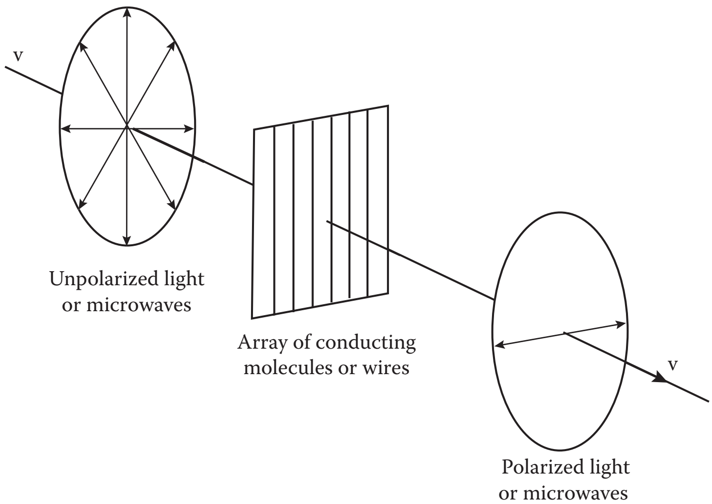


FIGURE 21.16 An analogy between polarized light and polarized microwaves.

Indeed, there exists a special angle (called the *Brewster angle*) where the reflected light wave is totally polarized (Fresnel's equations) and the refracted wave is partially polarized (see Figure 21.17). The reflected beam will be completely polarized when the angle between the reflected and refracted rays is 90°. The incident angle at which this effect occurs is the Brewster angle (θ_B), named after a Scotsman, Sir David Brewster (1781–1868).

To determine θ_B, use Snell's law of refraction.

$$n_1 \sin \theta_2 = n_2 \sin \theta_2.$$

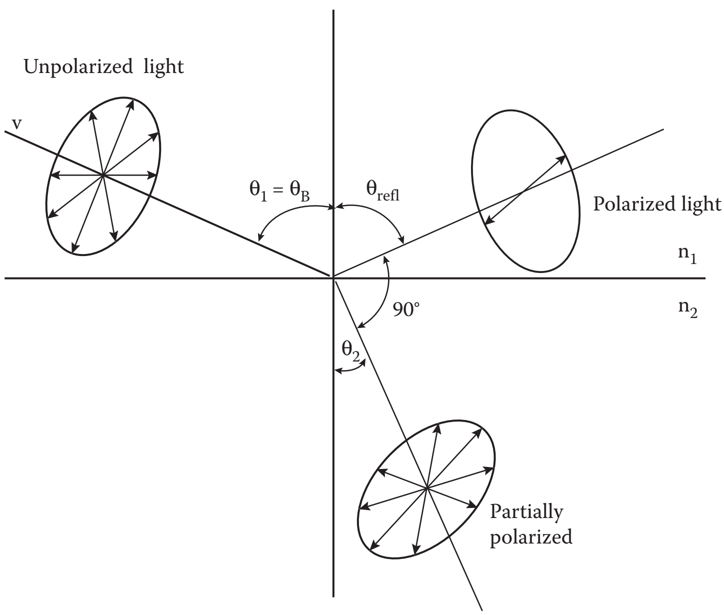


FIGURE 21.17 Polarization by reflection, the Brewster angle.

From Figure 21.17, $\theta_B = \theta_{refl}$, so $\theta_B + \theta_2 = 90°$, and Equation 21.22 can be written as

$$n_1 \sin \theta_B = n_2 \sin(90 - \theta_B) = n_2(\sin 90° \cos \theta_B - \sin \theta_B \cos 90)$$

or

$$n_1 \sin \theta_B = n_2 \cos \theta_B.$$

Finally,

$$\frac{\sin \theta_B}{\cos \theta_B} = \tan \theta_B = \left(\frac{n_2}{n_1}\right). \tag{21.23}$$

Equation 21.23 is known as *Brewster's law.*

EXAMPLE 21.8

Determine the Brewster angle for light traveling in air and incident on the smooth surface of a "calm" swimming pool.

Solution

$$n_1 = n_{air} = 1.000 \text{ and } n_2 = n_{water} = 1.333.$$

Brewster's law gives

$$\text{Tan } \theta_B = \left(\frac{n_2}{n_1}\right) = \left(\frac{1.333}{1.000}\right) = 1.333$$

or

$$\theta_B = \tan^{-1}(1.333) = 53.1°.$$

Double Refraction: Calcite and other transparent, anisotropic crystals display unusual refraction properties. Unpolarized light, incident in certain directions, on such crystals, results in two, not one, refracted rays. One refracted ray obeys Snell's law of refraction and is appropriately called the *ordinary ray*. The other refracted ray, called the *extraordinary ray*, does not obey Snell's law and often has a direction that is not even in the incident plane. This double refraction is called *birefringence*. Discussion of its microscopic cause is beyond the scope of this text. The most important property of this birefringence is that the two refracted rays are totally plane polarized in two mutually perpendicular directions. So, the ordinary ray polarization direction is perpendicular to the extraordinary ray polarization direction. Since these two rays propagate in different directions, they become spatially separated and can be analyzed independently.

Other properties of these rays are as follows:

1. Equal amounts of the incident unpolarized light go into the two polarized rays. That is, each receives half of the incident light. Contrast this with reflection where a small amount of the incident beam, much less than half, is polarized at the Brewster angle.
2. Only a small amount of the incident light is absorbed by the transparent birefringent crystal. This polarization process is very energy efficient.

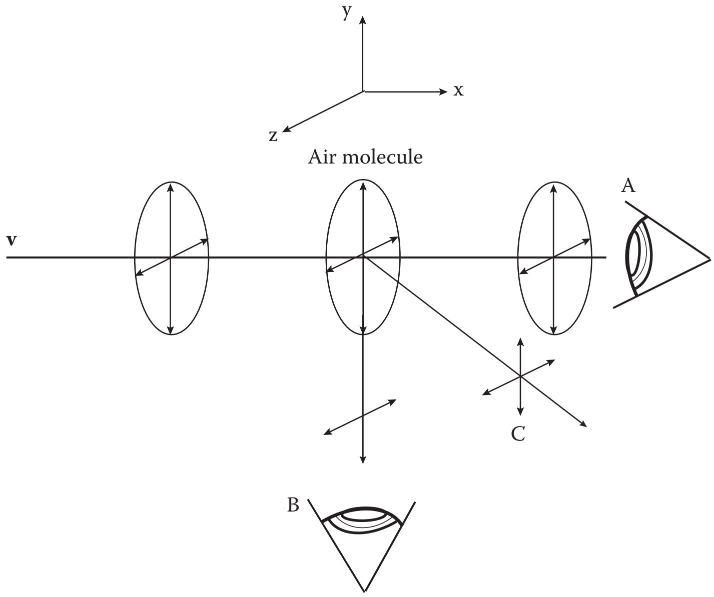


FIGURE 21.18 Scattering from air molecules partially polarizes sunlight.

Scattering: When light encounters an isolated atom/molecule, some of the light energy is absorbed by the atom/molecules electrons. This causes them to oscillate about the center of positive charge with the same frequency as the incident light. These vibrating electrons behave as small dipole antennae and reradiate light of the same frequency. None of the reradiated light is along the axis of vibration. Some of the reradiated light combines with the incident light to form new wavefronts (LSP). This process can occur with one atom/molecule or with many as when light is transmitted through a gas or a transparent liquid or solid. The process is called *scattering*. Scattering partially polarizes the sunlight that passes through our atmosphere. The light that reaches the earth's surface is a combination of direct sunlight and reradiated light. Because the reradiated light is not emitted along the axis of electron vibration, but is emitted perpendicular to it, the resulting light is partially polarized. Figure 21.18 shows unpolarized light from the sun incident on an atmospheric molecule. The many components of **E** have been resolved into the two total components in the y–z plane. Similarly, the induced vibrations of the molecules electrons can be perceived as oscillating along the y and z axes.

When viewed along the x axis, that is, perpendicular to the electrons plane of oscillation (position A), the reradiated light is unpolarized. The view along the y axis is in a direction in which the reradiating molecule emits no light energy (position B). Thus, only the components of **E** that are parallel to the z axis are seen. The light is polarized. When viewed along intermediate locations, the light is partially polarized (position C).

PROBLEMS

21.1 What wavelength (in nm) of a monochromatic source in the visible region can be used to constructively reflect off a soap film ($n = 1.46$) if the latter is 210 nm thick? *Note:* $1.0\text{ nm} = 10^{-9}\text{ m}$.

21.2 A double-slit arrangement produces a second-order constructive interference fringe 2.30° from the central maximum when irradiated with light of wavelength $\lambda = 589$ nm. Determine the slit separation "d."

21.3 Light of wavelength $\lambda = 5.00 \times 10^{-7}$ m is incident on a single slit of width "a." The third minimum in the diffraction pattern occurs at an angle of 1.50° (measured from the central maximum). Determine the slit width "a."

21.4 A Young's double-slit apparatus has a slit separation of $d = 1.60 \times 10^{-4}$ m and a viewing screen positioned 3.00 m from the slits. What value of wavelength is required for the separation of adjacent fringes to be displayed 1.20 cm apart on the screen.

21.5 A thin film has an index of refraction $n = 1.45$. It exhibits constructive interference for light of wavelength 4.30×10^{-7} m and destructive interference for light of wavelength 6.65×10^{-7} m. Determine the thickness of the film.

21.6 Determine the incident angle (Brewster's angle) for reflected light to be completely polarized when traveling from air to (a) acrylic plastic, $n_{ap} = 1.510$, and (b) diamond, $n_D = 2.417$.

21.7 A radio station broadcasts at a carrier frequency of 640.0 kHz. The station broadcasts in-phase signals from two antennas, one that is situated 4925 km due west of your home, the other is 8442 km due east. Does constructive or destructive interference occur at your home receiving antenna? Establish your answer.

21.8 A camera lens ($n_l = 1.54$) is to be covered with a nonreflective coating of magnesium fluoride ($n_c = 1.38$) to prevent reflections of green light ($\lambda = 550$ nm) from the lens. Determine the minimum, nonzero coating thickness required.

21.9 A soap film ($n_s = 1.35$) is illuminated perpendicularly by 600.0 nm light. The film is immersed in air. For what minimum value film thickness does constructive interference occur?

21.10 Light of wavelength $\lambda = 610.0$ nm is directed onto two narrow slits separated by 0.157 mm. The interference pattern is observed on a flat screen placed at $L = 1.74$ m from the slits. Determine (a) the angular location of the first-order dark fringe and (b) the linear width, along the screen, from the central bright fringe.

21.11 Light of wavelength $\lambda = 650.0$ nm is directed toward a single slit and forms a diffraction pattern. Determine the angular location of the first dark fringe for slits of width: (a) 1.710×10^{-6} m and (b) 1.710×10^{-4} m.

21.12 Diffraction by a large "slit." A standard inside doorway is 32.0 inches = 0.813 m wide. Determine (a) the angular location of the first dark fringe if light of wavelength $\lambda = 450.0$ nm is directed toward the doorway, and (b) the first intensity minimum of sound waves with frequency $f = 459.0$ Hz and velocity $v = 345.0$ m/s that pass through the doorway.

21.13 Light of wavelength $\lambda = 550.0$ nm is incident on a diffraction grating. A second-order maximum is observed at the angular location of 16.0°. Determine the number of lines/cm for this grating.

21.14 Light of wavelength $\lambda = 560.0$ nm is incident on a diffraction grating with 1.13×10^5 lines/cm. Calculate the angular location of the third order maximum for this light.

21.15 A beam of plane-polarized light is incident on a sheet of Polaroid. At what angle θ should the transmission axis of the Polaroid be oriented so that the intensity of the transmitted beam is reduced to three-fourths of its incident value?

22 Modern Physics

22.1 INTRODUCTION

In the beginning of the twentieth century, there occurred a surge in both the experimental data acquired and the theories proposed to explain the sometimes "unexpected" data. With the formulation of the Special Theory of Relativity by Albert Einstein in 1905, the "quantum hypothesis" by Planck in 1900, and the development of quantum theory in the 1920s and 1930s, modern physics was created. This veritable explosion of data and theory led to an entirely new and different perception of the laws of nature and the way physical systems behave.

In this chapter, some of the modern physics data and theory are addressed. The material is not discussed chronologically but presented in an attempt to form a complete coherent explanation of physics behavior.

22.2 ATOMS

If a chemical element is subdivided into smaller and smaller quantities, it will ultimately be reduced to its "smallest" amount that still retains the chemical properties of the element. That smallest amount is called an *atom*. Indeed, the word atom is based on the Greek word *atomos*, which means indivisible.

In 1911, the New Zealand physicist Ernest Rutherford (1871–1937) and coresearchers did a series of experiments whereby they directed a beam of alpha (α) particles (doubly ionized helium atoms) at a thin gold foil. The scattering pattern of these α particles (some scattered completely backward) led Rutherford to conclude that the gold consisted of particles with large, positively charged centers (nuclei). Experiments by the English scientist J.J. Thomson (1856–1940) on the charge to mass ratio of cathode rays led to the concept of free electrons. The existence of light (low mass), negatively charged electrons, and heavier, positively charged protons led to planetary model (Rutherford model) of the atom. From these experiments, it was concluded that an atom consisted of a relatively massive positive nucleus surrounded by less massive, negatively charged electrons revolving around the nucleus (Figure 22.1).

The planetary model resembles our solar system with the less massive planets revolving around a central, more massive sun. Later measurements determined atomic diameters to be approximately 10^{-10} m and nuclear diameters to be approximately 10^{-15} m. This means that an atom consists of mostly empty space.

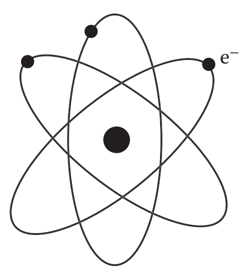


FIGURE 22.1 The planetary (Rutherford) model of an atom.

22.3 LINE SPECTRA

In Chapter 20, it was stated that electromagnetic waves (EMWs) can be created either by accelerating charged particles or by electron or nucleon energy transitions in atoms. If the atoms of a particular chemical element are relatively free, they can be made to emit EMWs whose wavelengths are characteristics of that particular element. This can be achieved by putting a low-pressure gas in a sealed transparent tube equipped with electrical conduction electrodes, placed within the tube, at each end. Applying a large potential difference (electric field) to the electrodes causes the atom to emit EMWs. By using a diffraction grating spectrometer, described in Chapter 21, the individual wavelengths emitted can be separated into bright fringes and numerical values determined. The group of fringes from a particular element is called its *line spectrum*. Not all the spectral lines from a given element are in the visible range of EMWs. The simplest atom, and the one with the simplest line spectrum, is atomic hydrogen. Its visible line spectrum was investigated by the Swiss schoolteacher Johann Balmer (1825–1898) and is appropriately called the *Balmer series*. Balmer constructed a generalized empirical equation that gives the values of the observed wavelengths. His equation is

$$\frac{1}{\lambda} = R\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right), \quad n_i = 2, n_f = 3, 4, 5. \tag{22.1}$$

Here n_i and n_f are positive integers with $n_f > n_i$ and $R = 1.097 \times 10^7\ m^{-1}$.

Two other spectral series are described by equations of the same form as Equation 22.1. They are, the *Lyman series* ($n_i = 1$, $n_f > n_i$) that applies to shorter wavelengths and the *Paschen series* ($n_i = 3$, $n_f > n_i$) that predicts longer wavelength values.

Equation 22.1 and the equations for the other two spectral series make it possible to catalog the spectra of the gaseous state of many elements. Then, if compared with the catalog of spectra, an originally unknown gas can be identified. Each element has its own characteristic line spectrum, analogous to humans each having a unique fingerprint.

Equation 22.1 can also be used to determine the wavelength limits of a series, as shown in the following example.

EXAMPLE 22.1

Determine the wavelength limits of the Balmer series.

SOLUTION

Each integer n_f corresponds to a given wavelength in the series and the larger is n_f, the smaller is λ. For the shortest wavelength, $n_f = \infty$,

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{\infty}\right) = R\left(\frac{1}{4} - 0\right) = \frac{R}{4}.$$

So,

$$\lambda = \frac{4}{R} = \frac{4}{(1.097 \times 10^7\,m^{-1})} = 3.65 \times 10^{-7}\,m = 365nm.$$

For the longest wavelength, $n_f = 3$,

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{3^2}\right) = R\left(\frac{1}{4} - \frac{1}{9}\right) = R\left(\frac{5}{36}\right)$$

or

$$\lambda = \frac{36}{5R} = \frac{36}{5(1.097 \times 10^7\,\text{m}^{-1})} = 6.56 \times 10^{-7}\,\text{m} = 656\,\text{nm}.$$

Equation 22.1 gives the atomic hydrogen line spectrum but does not reveal *why* only certain wavelengths are emitted. The Danish physicist Niels Bohr (1885–1962) constructed the first model of atomic hydrogen that predicted its discrete spectral fringes.

22.4 BOHR MODEL

To construct his model, Bohr used the Rutherford, classical planetary model of the atom and incorporated some new quantum ideas from Planck and Einstein. Einstein, and others before him, assumed that electromagnetic radiation could be considered as being made up of either waves or particles. The particles are called *photons*, so a beam of light could be perceived as a stream of photons. Planck assumed that these photons had quantized energy values. That is, they could have only specific values of energy given by

$$E = hf. \tag{22.2}$$

Here, h is a Planck's constant and has the value $h = 6.626 \times 10^{-34}$ J s and f is the frequency of the light.

EXAMPLE 22.2

Consider a 60.0-W light source (assumed to be 100% efficient so that all the power is emitted as light, none as heat emits light (photons) of wavelength $\lambda = 550.0$ nm. Determine the number of photons it emits per second.

Solution

The frequency is acquired from the wavelength by $f = c/\lambda$. If each photon has energy $E_1 = hf$, n photons have energy $E_{total} = nhf$. So,

$$\text{Power} = \frac{\text{energy}}{\text{time}} = \frac{nhf}{t} = \frac{nhc}{t\lambda} = P$$

or

$$\left(\frac{n}{t}\right) = \frac{P\lambda}{hc} = \frac{(60.0\,\text{W})(5.500 \times 10^{-7}\,\text{m})}{(6.63 \times 10^{-34}\,\text{J s})(3.00 \times 10^8\,\text{m/s})} = 1.66 \times 10^{20}\,\frac{\text{photons}}{\text{second}}.$$

To develop his model, Bohr assumed that the mechanical energy of an atom is quantized, that is, exists in discrete quantities just as earlier, Planck had assumed that the energy of EMWs is quantized. Addressing the simplest atom, hydrogen, Bohr made some other assumptions. They were (a) the planetary model is valid, (b) the electron orbits are circular and the only ones permitted are those whereby the orbital angular momentum is quantized according to $\ell = (nh)/(2\pi)$, $n = 0,1,2,3,\ldots$, (c) the centripetally accelerating orbital electrons, while in their allowed orbits, do not radiate energy, and (d) orbital electrons can make transitions from one allowed orbit to another, either by the atom absorbing or emitting a photon. That is, if E_i represents the energy of the electron in its initial orbit and E_f its energy in the final orbit, then

$$\pm(E_f - E_i) = hf. \tag{22.3}$$

If $E_f > E_i$, the electron goes from a lower energy to a higher energy by absorbing a photon of frequency f. In that case, the (+) sign in Equation 22.3 pertains. If $E_f < E_i$, the electron (atom) looses energy and emits a photon. In that case, the (–) sign is appropriate (see Figure 22.2).

Using the above assumptions, and recognizing that the allowed electron orbits are due to their Coulomb attraction to the positive nucleus, Bohr was able to derive an expression for the possible atomic hydrogen energy values E_n. He found:

$$E_n = -K\left(\frac{Z^2}{n^2}\right), \quad n = 1,2,3,.... \tag{22.4}$$

Here, Z is the atomic charge number and Z = 1 for atomic hydrogen. The constant K is given by

$$K = \left(\frac{2\pi^2 m_e k^2 e^4}{h^2}\right). \tag{22.5}$$

Numerically, $K = 2.18 \times 10^{-18}$ J.

The negative sign in Equation 22.4 indicates that the electrons are in *bound* energy states. Note that as n increases, E_n becomes a smaller negative value, that is, E_n also increases.

Bohr also derived an expression for the radius of the nth orbit (*Bohr orbit*) as

$$r_n = \left(\frac{h^2}{4\pi^2 m_e k e^2}\right)\frac{n^2}{Z} = (5.29 \times 10^{-11}\ \text{m})\left(\frac{n^2}{Z}\right). \tag{22.6}$$

For the hydrogen atom (Z = 1), the smallest Bohr orbit (n = 1) is called the *Bohr radius* and has the value

$$r_1 = 5.29 \times 10^{-11}\ \text{m}.$$

By inserting Equation 22.4 into Equation 22.3, Bohr was able to predict the line spectrum of atomic hydrogen. Note that for *emission* spectra, $E_i > E_f$, so Equation 22.3 is

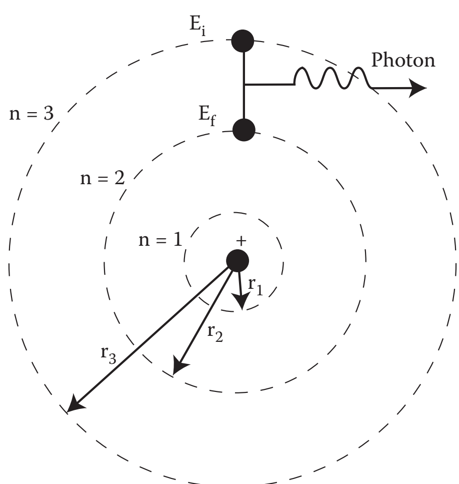


FIGURE 22.2 Atomic electrons can make transitions between discrete energy levels by absorbing or emitting a photon.

$$(E_i - E_f) = hf = \frac{hc}{\lambda}.$$

Then,

$$-KZ^2\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right) = \frac{hc}{\lambda}$$

or

$$\frac{1}{\lambda} = \frac{KZ^2}{hc}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right), \quad n_i > n_f,\ n_i, n_f = 1, 2, 3, 4, \ldots \tag{22.7}$$

Numerically,

$$\frac{1}{\lambda} = (1.097 \times 10^7\,\mathrm{m}^{-1})\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right). \tag{22.8}$$

Bohr's Equation 22.8 is identical to Balmer's Equation 22.1.

EXAMPLE 22.3

Determine the wavelength of light emitted from gaseous hydrogen when electrons in the $n = 4$ state (orbit) make transitions to the $n = 2$ state.

Solution

Using Equation 22.7 with $n_i = 4$ and $n_f = 2$,

$$\frac{1}{\lambda} = (1.097 \times 10^7\,\mathrm{m}^{-1})\left(\frac{1}{2^2} - \frac{1}{4^2}\right) = R\left(\frac{1}{4} - \frac{1}{16}\right) = R\left(\frac{4-1}{16}\right) = \frac{3R}{16}$$

or

$$\lambda = \frac{16}{3R} = \frac{16}{3(1.097 \times 10^7\,\mathrm{m}^{-1})} = 4.86 \times 10^{-7}\,\mathrm{m} = 486\,\mathrm{nm}.$$

This is a blue-green color.

Equation 22.4 can be used to determine the ground-state energy and ionization energy of atomic hydrogen. It is convenient to express these energies in units of electron volts (eV). Recall that (1.0) eV is the energy acquired by an electron in moving through a potential difference of 1 V. So,

$$(1.0)\ \mathrm{eV} = (1.0)(1.60 \times 10^{-19}\,\mathrm{C})(1.0\ \mathrm{J/C}) = 1.60 \times 10^{-19}\ \mathrm{J}.$$

The ground state of hydrogen ($n = 1$) from Equation 22.4 is

$$E_{n=1} = -K\left(\frac{Z^2}{n^2}\right) = -(2.18 \times 10^{-18}\,\mathrm{J})\left(\frac{1.0}{(1.0)^2}\right) = -(2.18 \times 10^{-18}\,\mathrm{J}).$$

Expressing in eV,

$$E_{n=1} = -\frac{(2.18 \times 10^{-18}\text{ J})}{(1.60 \times 10^{-19}\text{ J/eV})} = -13.6\text{ eV}. \tag{22.9}$$

So, the ground-state electron is bound to the nucleus with an energy of 13.6 eV. This is also the amount of energy required to strip that electron free from the nucleus, that is, the *ionization energy*.

So, Equation 22.4 can equivalently be written as

$$E_n = (-13.6\text{ eV})\left(\frac{Z^2}{n^2}\right). \tag{22.4a}$$

22.5 BEYOND THE BOHR MODEL

Although the Bohr model was successful in approximately predicting the line spectrum of atomic hydrogen, application of it to multielectron atoms was unsuccessful. Later, high-resolution spectrometers revealed that spectral lines were really two separate, closely spaced lines, called *fine structure*. These close splittings led some to believe that the simple Bohr picture of electron transitions was incomplete. It became evident that some other type of motion or energy was responsible for the splittings and the more complicated line spectra of multielectron atoms. Additional theoretical work by German physicist Arnold Sommerfeld (1868–1951), Austrian physicist Wolfgang Pauli (1900–1958), and Dutch physicists Samuel Goudsmit and George Uhlenbeck resulted in a scheme that required four numbers, called *quantum numbers*, to describe the behavior of electrons in atoms. Their contributions, essentially led to (i) assuming elliptical, as opposed to circular, orbits, (ii) including the magnetic energy of the atom, and (iii) postulating the existence of electron spin.

The inclusion of elliptical orbits led to two quantum numbers n_r and n_θ. These are usually combined as $n = (n_r + n_\theta)$ and called the *principal quantum number* n, similar to values in the Bohr model, that is,

$$n = 1, 2, 3, \ldots. \tag{22.10}$$

The orbital angular momentum (**L**), as with the Bohr model, was assumed to be quantized. Its magnitude was determined to be

$$L = \sqrt{\ell(\ell + 1)}\left(\frac{h}{2\pi}\right), \tag{22.11}$$

where ℓ, the *orbital quantum number*, is always positive and has the possible values,

$$\ell = 0, 1, 2, \ldots (n - 1). \tag{22.12}$$

Note that for the case $\ell = 0$, also $|L| = 0$, so that electron is not revolving around the nucleus, but is oscillating through it.

In addition, the allowed electron orbits may not all be in the same plane, but have different spatial orientations. To distinguish these orientations, the *magnetic quantum number* (m_ℓ) is introduced. It gives the *projection* of **L** along an arbitrarily chosen axis or direction, that is usually designated as the z axis. The possible values of m_l are

$$m_\ell = -\ell, -(\ell - 1), -(\ell - 2), \ldots, -2, -1, 0, +1, +2, \ldots, +\ell. \tag{22.13}$$

The z component of **L** is

$$L_z = m_\ell\left(\frac{h}{2\pi}\right). \tag{22.14}$$

The word "magnetic" is used to describe m_ℓ because an externally applied magnetic field affects the energy of the atom. The amount of influence is specified by the m_l value. If the atom is not in a $\mathbf{B}_{ext}$, m_ℓ has no effect on its energy.

To account for the two closely spaced spectral splittings, Goudsmit and Uhlenbeck postulated that atomic electrons have intrinsic spin, called *electron spin*, about an axis through, and moving with, the electron. It is convenient to view the electron as a spherical, spinning charged object made up of infinitesimal rings of charge rotating in planes perpendicular to the direction of the axis. As such, these rings constitute current loops and give rise to a magnetic moment $\boldsymbol{\mu}$. Thus, the spinning electron behaves as a tiny bar magnet. Recall from Equation 16.14 that a bar magnet, or a current loop, immersed in an external magnetic field has a potential energy given by

$$U = -|\mu||B_{ext}|\cos\theta. \tag{16.14}$$

The electron revolves about the positive nucleus, but from the frame of reference of the electron, the nucleus is revolving about the electron. This revolving nucleus constitutes an electric current and creates a **B** field at the location of the electron. Thus, the spinning electron "sees" a magnetic field due to its orbital motion and Equation 16.14 applies. Since the fine structure consists of two closely spaced lines, Goudsmit and Uhlenbeck reasoned that only two spatial orientations would satisfy Equation 16.14. They obtained an expression for the spin magnetic moment and calculated the *spin angular momentum* (**S**) to be

$$|S| = \frac{1}{2}\left(\frac{h}{2\pi}\right)$$

The two spatial orientations should thus be when the spin magnetic moment (μ_s) is either parallel or antiparallel to the orbital magnetic moment (μ_l). Another way of viewing this is that the axis of electron spin is perpendicular to the plane of its orbital motion. If μ_s is parallel to μ_l, this is called "spin up," if antiparallel it is "spin down." Thus, the z component of spin angular momentum is

$$S_z = \pm\frac{1}{2}\left(\frac{h}{2\pi}\right). \tag{22.15}$$

The *spin magnetic quantum number* (m_s) therefore has two possible values, that is,

$$m_s = +\frac{1}{2}, -\frac{1}{2}. \tag{22.16}$$

The improvement to the Bohr model now requires four quantum numbers: n, l, m_ℓ, and m_s, to describe each state of the hydrogen atom. One set of values corresponds to one given state.

EXAMPLE 22.4

Determine the number of possible states for the hydrogen atom when $n = 2$.

Solution

From Equation 22.12, the largest value of $\ell = (n - 1) = (2 - 1) = 1$, so ℓ can be equal to 0 or 1. Use Equations 22.13 and 22.16 to determine possible allowed values for m_ℓ and m_s.

$$\text{For } \ell = 0,\ m_\ell = 0 \text{ and } m_s = +(1/2) \text{ or } m_s = -(1/2).$$

For $\ell = 1$, $m_\ell = -1, 0, +1$ and $m_s = +(1/2)$ or $m_s = -(1/2)$ for each value of m_ℓ. It is instructive to construct a table of possible values and states (Table 22.1).

For $n = 2$, there are eight possible states all with the same energy.

$$E_n = (-13.6\text{eV})\left(\frac{Z^2}{n^2}\right) = (-13.6\text{ eV})\left(\frac{(1.0)^2}{(2^2)}\right) = -3.4\text{ eV}.$$

The restrictions on the values of n, l, m_ℓ, and m_s, that is, Equations 22.10, 22.12, 22.13, and 22.16, respectively, are the result of a theoretical assumption. It is that the number of electrons that can occupy a given energy state is limited. This restriction is summarized in the *Pauli exclusion principle*, which states: *No two electrons in an atom can have the exact same numerical values for the quantum numbers n, ℓ, m_ℓ, and m_s.*

In the previous example, these restrictions can be observed by comparing the eight columns in Table 22.1.

22.6 PERIODIC TABLE OF ELEMENTS

The Russian chemist Dmitri Mendeleev (1834–1907) developed the original periodic table by grouping the known elements according to the similarities of their chemical properties. These properties are due to the particular elements outer electron configuration. As will be shown, quantum theory and the Pauli exclusion principle can predict these configurations.

With the exception of atomic hydrogen, all neutral, that is, nonionized, atoms have more than one electron. These multielectrons have, in addition to their Coulomb attraction to their nucleus, a Coulomb repulsion for each other. This repulsion contributes to the total energy of the atom. Thus, the single electron energy expression, Equation 22.4, that is,

$$E_n = (-13.6\text{ eV})\left(\frac{Z^2}{n^2}\right),$$

does not apply to multielectron atoms. Nevertheless, the conventional approach to treating these atoms is via the use of the four quantum numbers n, ℓ, m_ℓ, and m_s.

Quantum mechanical calculations show that the energy level of each state depends not only on the principle quantum number n, but also on the orbital quantum number ℓ. The energy of these

TABLE 22.1
Quantum Numbers

n	2	2	2	2	2	2	2	2
ℓ	0	0	1	1	1	1	1	1
m_ℓ	0	0	−1	−1	0	0	+1	+1
m_s	+(1/2)	−(1/2)	+(1/2)	−(1/2)	+(1/2)	−(1/2)	+(1/2)	⊠(1/2)

states is designated $E_{n,\ell}$. Generally $E_{n,\ell}$ increases as n increases, but there are some exceptions to this pattern. Additionally, for a specific n, $E_{n\ell}$ increases as ℓ increases. These levels are designated by horizontal lines as shown in Figure 22.3.

The lowest energy state of an atom is called the *ground state*. As is true in most physical systems, the electrons in atoms tend to seek or occupy the lowest energy state. At room temperature and lower temperature, the electrons in atoms spend most of their time in the ground state.

When a multielectron atom is in its ground state, not all of the electrons are in the n = 1 level because of the Pauli exclusion principle. A note on nomenclature: electrons with the same value of n are in the same *shell*. Those with the same values of *both* n and l are said to be in the same *subshell*. For example, the n = 3 shell has three subshells, one each for

$$\ell = 0, 1, \text{ and } 2.$$

The periodic table of elements can be constructed or "built up" by filling the shells and subshells, from the lowest energy to the highest, that is, from the bottom up, in compliance with the Pauli exclusion principle.

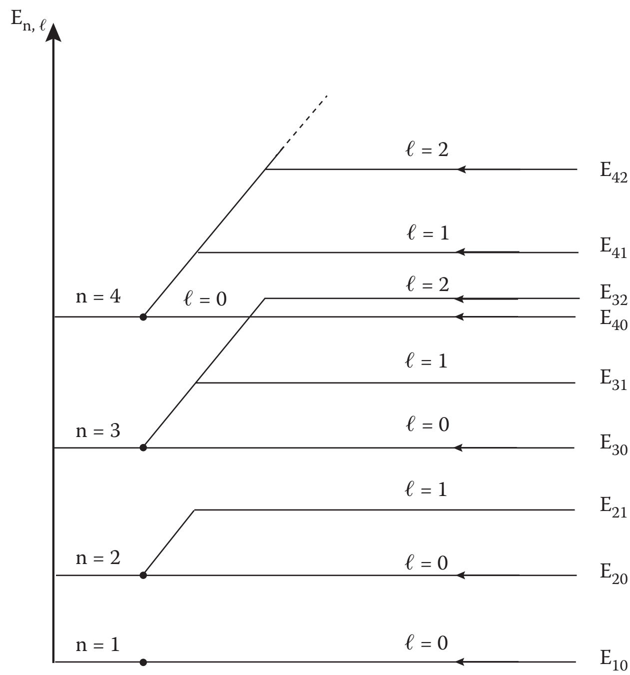


FIGURE 22.3 The atomic energy levels are represented by horizontal lines.

22.6.1 Ground States of Atoms

The ground state, occupied levels of the first eight elements of the periodic table, are shown in Figure 22.4. The dots represent electrons.

Subshell n = 1, $\ell = 0$ has the lowest value of energy and for it, $m_\ell = 0$ and $m_s = \pm(1/2)$. Since atomic hydrogen has only one electron, its spin value can be either (+1/2) or (–1/2). All the other atoms have two or more electrons, so in this subshell, one electron will have $m_s = (+1/2)$ and the other $m_s = (-1/2)$. Shell n = 2 has two subshells; n = 2, $\ell = 0$, and n = 2, $\ell = 1$. For $\ell = 0$, $m_\ell = 0$, and for $\ell = 1$, $m_\ell = -1, 0, +1$. The first subshell (n = 2, $\ell = 0$) can accommodate two electrons, the second subshell (n = 2, $\ell = 1$), six electrons, that is, the two m_s values for each of the three m_ℓ values. So

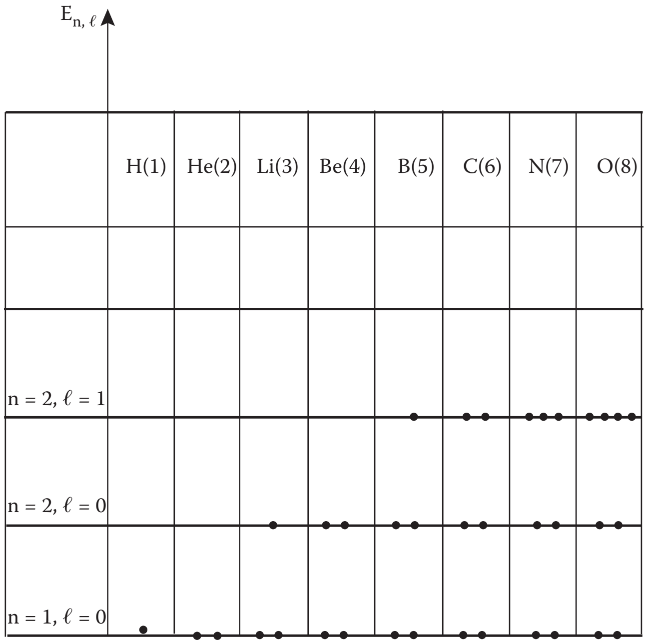


FIGURE 22.4 The ground-state, occupied energy levels of some elements.

eight electrons, two plus six can occupy shell n = 2, and so on. In this manner, the electron occupancy of energy levels can be determined.

EXAMPLE 22.5

Determine which of the possible energy levels of sodium (Na) its 11 electrons occupy in the ground state.

Solution

Shell $n = 1$, $\ell = 0$ has $m_\ell = 0$, $m_s = \pm(1/2)$ holds two electrons.
Shell $n = 2$, $\ell = 0$, $m_\ell = 0$, $m_s = \pm(1/2)$ holds two more electrons.
Shell $n = 2$, $\ell = 1$, $m_\ell = 0$, $m_s = \pm(1/2)$ is occupied by six electrons.
Shell $n = 3$, $\ell = 0$, $m_\ell = 0$, $m_s = \pm(1/2)$ can hold two electrons, but is occupied by only one electron since the first three subshells already contain 10 of sodium's 11 electrons.

22.7 X-RAYS

An x-ray spectrum can be produced by bombarding a heavy metal target, usually platinum or molybdenum, with a beam of high-energy electrons as shown in Figure 22.5.

The spectrum (Figure 22.6) can be explained as a combination of radiation resulting from both accelerating charged particles (electrons) as mentioned in Chapter 19 and electron energy transitions in atoms.

The continuous part of the spectrum results from the deceleration (negative acceleration) of the electrons in the electron beam. They impinge on the target and their velocity is decreased from a large value to zero in a very short time interval. All the beam electrons are not stopped simultaneously since some penetrate more deeply into the target than others. Thus, their stopping times vary, yielding varying accelerations. This gives a continuous distribution of EMW in the x-ray region. Indeed, the continuous part of the spectrum is called *bremsstrahlung* (in German for "braking" radiation or stopping radiation). The peaks or "spikes" in Figure 22.6 result from electron transitions

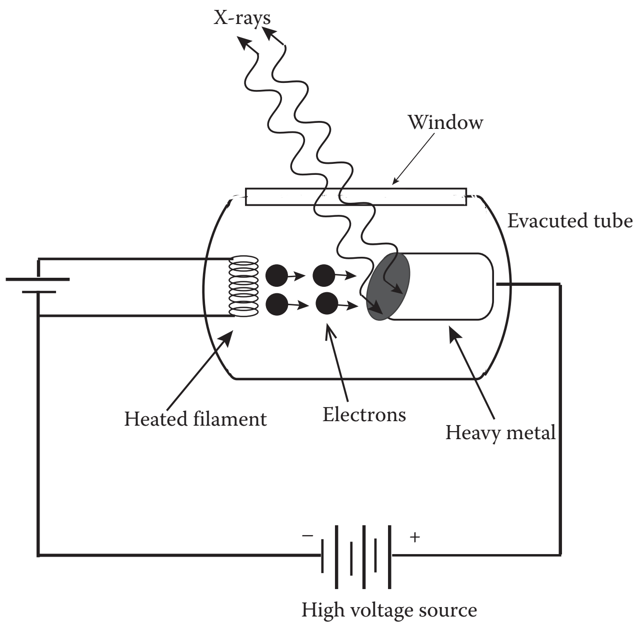


FIGURE 22.5 High-energy electrons, colliding into a metal target, can produce x-rays.

into the n = 1 shell of the metal target. Chemists call the n = 1, the "K" shell, so the peaks in Figure 22.6 are labeled K_α and K_β. The K_α peak results from electron transitions, in the target atoms, from shell n = 2 to n = 1. The K_β results from transitions from n = 3 to n = 1. These transitions are possible when a beam electron with enough energy hits a target atom and knocks an n = 1 shell electron out of the atom. An electron from one of the atoms outer shells can then fall into the n = 1 shell and emit an x-ray photon. The energy difference (ΔE) between the n = 3 and n = 1 states, that is, ($\Delta E_{3,1}$) is greater than ($\Delta E_{2,1}$) so the K_β x-rays have a higher energy and lower wavelength than the K_α x-rays.

The description of x-rays is a good example of the accurate predictions of *both* classical electrodynamics and quantum theory.

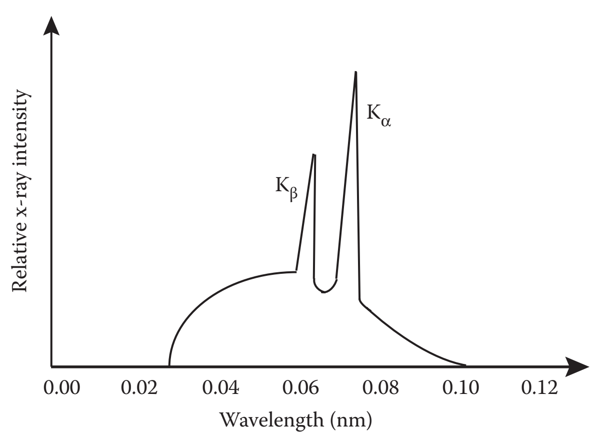


FIGURE 22.6 The x-ray spectrum consists of both continuous and discrete regions.

EXAMPLE 22.6

Estimate, via the Bohr model, (a) the minimum energy required to remove an n = 1 electron from a molybdenum (Z = 42) target and (b) the wavelength of the K_β (n = 3 to n = 1) x-rays.

Solution

Note 1: The Bohr model, Equation 22.4, applies to single electron atoms or ions, but is sometimes used to make approximations to many electron atoms.

Note 2: Using Equation 22.4

$$E_n = (-13.6\text{ eV})\left(\frac{Z^2}{n^2}\right).$$

To approximate the energy of the n = 1 electron, the quantity Z should be replaced by (Z – 1) to account for the "shielding," from the nucleus of one electron by the other.

a. To remove an n = 1 electron,

$$E_n = (-13.6\text{ eV})\left(\frac{Z^2}{n^2}\right) \to (-13.6\text{ eV})\left(\frac{(Z-1)^2}{n^2}\right) = (-13.6\text{ eV})\frac{(41)^2}{(1^2)} = -2.29 \times 10^4\text{ eV}.$$

So, the minimum KE of the beam electrons must be 2.29×10^4 eV. Thus, the beam electrons must be accelerated through a voltage of at least 22,900 V.

b. Using Equation 22.7

$$\frac{1}{\lambda} = (1.097 \times 10^7\,\text{m}^{-1})(Z^2)\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) = (1.097 \times 10^7\,\text{m}^{-1})(41)^2\left(\frac{1}{1^2} - \frac{1}{3^2}\right)$$

or

$$\lambda = \frac{9}{8(41)^2\,(1.097 \times 10^7\,\text{m}^{-1})} = 6.10 \times 10^{-11}\text{m}.$$

Some of the properties of x-rays are as follows:

1. They affect photographic film, so film may be used to detect them.
2. Material objects have varying transparency to x-rays. They pass through glass, wood, and flesh rather easily, but are absorbed more by bone and metals.

Because of these properties, x-rays may be used as a medical diagnostic tool to detect broken bones, tumors, and so on. Because of their wave nature, they may also be used to study the structure of crystals and other solids.

22.8 QUANTUM PHYSICS

Our knowledge and understanding of the nature of the inanimate world has grown and evolved for some more than 5000 years. For a long time, parts of that knowledge seemed isolated and disconnected. On rare occasions, an individual or person assesses that large body of information has a flash of insight and connects the apparently disjointed knowledge by a grand theory. Such a connection first occurred in the mid-seventeenth century by Sir Isaac Newton (1642–1727).

He produced a complete mathematical description of the motion of objects that are acted upon by known forces and summarized by his three laws of motion. This branch of physics is called *mechanics*, and Newton's laws completely revolutionized our understanding and predictive capabilities about motion.

A second revolution in thought occurred in 1865 when James Clerk Maxwell published his "Maxwell's Equations." These equations were a compilation of the experimental work of Gauss, Ampere, Faraday, and Maxwell. These equations describe all the phenomena of classical electricity and magnetism and indeed led to the prediction of wireless communications.

Between these two revolutions, and afterwards, a steady stream of insights into the nature of both optics and thermodynamics were acquired.

A third revolution occurred in 1905 when Albert Einstein published his paper on *Special Relativity*. This theory introduced the concept of space–time, predicted the interchangeability of mass and energy, and resolved the confusion regarding absolute reference frames.

The fourth revolution, *quantum theory*, occurred approximately during the years 1900–1930. The quantum revolution is not attributed to any single individual, but consists of the contributions and explanations of many scientists. These scientists used quantum ideas to explain the results of several critical experiments that could not be explained by the classical theories of Newton and Maxwell.

Quantum theory invokes the quantum hypothesis. A superficial, qualitative treatment of thermal radiation will here be presented to illustrate the ad hoc unorthodox flavor of the quantum hypothesis.

22.9 THERMAL RADIATION

Any material object at nonzero temperature and in thermal equilibrium with its surroundings continuously absorbs and emits EMWs from its surface. If the object is not in thermal equilibrium, but is at a higher temperature than its surroundings, it emits more EMW than it absorbs. If at a lower temperature than its surroundings, it absorbs more than it emits.

The energy per unit area per unit time, that is, the intensity R emitted or absorbed by an object, was determined, experimentally and later theoretically, to be proportional to the fourth power of its kelvin temperature. This relation is given by the Stefan–Boltzmann radiation law that is expressed by the equation

$$R = \sigma e(T_0^4 - T_s^4). \tag{22.17}$$

Here, T_0 is the objects temperature, T_s the temperature of the surroundings, and e the *emissivity* ($0 \le e \le 1$) and a measure of the efficiency with which an object emits thermal radiation. σ is the *Stefan's constant*, which is given by

$$\sigma = 5.6703 \times 10^{-8}\left(\frac{W}{m^2\ K^4}\right).$$

The temperatures in Equation 22.17 must be expressed in kelvins.

EXAMPLE 22.7

A person whose body temperature is $T_0 = 98.6°F = 37.0°C = 310$ K steps out of a shower into a cooler room with an air temperature of $T_s = 72.0°F = 22.2°C = 295$ K. Assume the unclothed human body has an emissivity of $e = 0.750$. Determine (a) the net power per unit area radiated by the person to his surroundings and (b) the total power he radiates to the room if his body surface area is 1.40 m^2.

Solution

a. Equation 22.7 gives

$$R = \sigma e(T_0^4 - T_s^4) = \left(5.67 \times 10^{-8}\ \frac{W}{m^2K^4}\right)(0.750)\left[(310)^4 - (295)^4\right] = 70.7\ \frac{W}{m^2}.$$

b.

$$R \to \frac{\text{power}}{\text{area}} = \frac{P}{A}.$$

So,

$$P = RA = \left(70.7\frac{W}{m^2}\right)(1.40\,m^2) = 99.0\frac{J}{s}.$$

Expressing in kcal,

$$P = \left(99.0\frac{J}{s}\right)\left(\frac{1.0\,\text{kcal}}{4186\,J}\right) = 2.36 \times 10^{-2}\frac{\text{kcal}}{s}.$$

Recall that 1.0 kcal equals 1.0 dietary Calories, so the man's body must furnish 0.0236 Calories every second to supply the radiated energy.

For an object to be in thermal equilibrium with its surroundings, it must emit as much thermal radiation as it absorbs. Indeed, a body that can absorb all the electromagnetic radiation impinging on it at any and all wavelengths and correspondingly emit EMWs freely at any and all wavelengths has an emissivity equal to one and is called a perfect *blackbody*. This raises the following question: at what wavelengths are these EMWs emitted? Early and later experiments show a distribution of emitted radiant energy that depends on the wavelength of the emitted EMWs, which in turn depends on the objects temperature (see Figure 22.7).

A nonburnable object heated to approximately 900 K will emit some of its radiation in the visible range. As objects are heated to temperatures above 900 K, they began to glow a dull red and then cherry red, orange, becoming white hot before melting. For example, the star Betelgeuse, in the constellation Orion, has a surface temperature of approximately 3000 K and its visible emissions

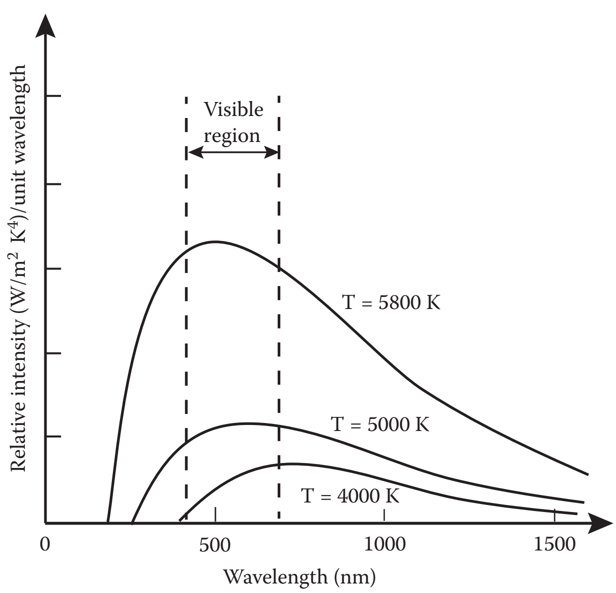


FIGURE 22.7 The maximum intensity of radiation emitted from an ideal blackbody depends on the blackbody temperature.

appear red-orange in color. Our sun, which is hotter with a surface temperature of approximately 5800 K, appears yellow. The human body, at 310 K, does not emit sufficient light so as to be seen in the dark with the unaided eye. It does emit sufficient EMWs in the infrared region, however, that can be measured with infrared detectors. Notice in Figure 22.7 that the wavelength at which the maximum emitted intensity occurs (say λ_{max}) shifts to smaller values as the temperature of the object increases. From such graphs, an empirical relation between λ_{max} and T was discovered by the German physicist Wilhelm Wien (1864–1928) and is appropriately known as the *Wien displacement law*. It states:

$$\lambda_{max} T = \text{constant} = 2.898 \times 10^{-3}\ \text{m K}. \tag{22.18}$$

This law implies that as the temperature increases, λ_{max} decreases, but their product remains constant for all temperatures. Notice also that the peak intensity of our sun is in the visible region of the EMW spectrum.

EXAMPLE 22.8

Use the Wien displacement law to determine λ_{max} for a human whose body temperature is $T = 310$ K.

Solution

$$\lambda_{max} = \frac{2.898 \times 10^{-3}\ \text{mK}}{310\ \text{K}} = 9.35 \times 10^{-6}\text{m} = 9350\ \text{nm}.$$

This value is in the infrared region of the EMW spectrum.

Efforts were made, applying the principles of the classical physics of Newton and Einstein to thermal radiation, that is, a blackbody radiator, to derive a mathematical fit to the data of Figure 22.7. These efforts were unsuccessful.

Much experimental study had been carried out on blackbody radiation, so the intensity-versus-wavelength data, similar to Figure 22.7, were considered valid. A good approximation to a blackbody radiator can be created in the laboratory by gouging a cavity into the interior of a large block of metal (Figure 22.8).

The imaginary surface of the small hole connecting the cavity to the outside wall of the metal block is treated as the surface of the blackbody. Thermal radiation inside the cavity is due to both radiation entering the cavity opening and the emissions due to the thermal vibrations of the atoms or molecules (wall oscillators) on the inside surface of the cavity. This interior radiation experiences many reflections and interactions with the wall oscillators. The cavity radiation field and the

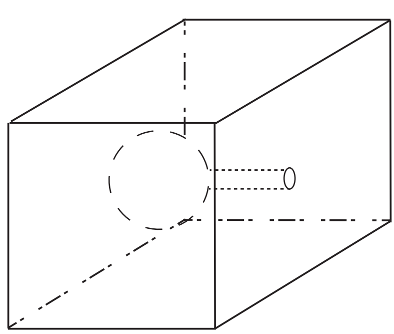


FIGURE 22.8 The small opening of the hole to a cavity in a block of metal behaves as a blackbody radiator.

oscillators exchange energy in thermal equilibrium and therefore mix all the modes (frequencies of the resultant standing waves). The radiation field can thus be thought of as a sum of frequencies via the linear superposition principle. So theoretically, analyzing the wall oscillators is equivalent to analyzing the radiation field. A small fraction of the cavity radiation escapes through the opening, which is treated as the blackbody surface.

In 1900, Max Planck (1858–1947), by introducing a strange and unorthodox quantum hypothesis, was able to replicate the thermal radiation curve for all wavelengths. Planck's analysis emphasized the wall oscillators, each with 2 degrees of freedom: their vibrational KE and PE. He assumed that the energies of these oscillators are restricted to integer multiples of a given energy, say E_0, for a particular oscillator mode of wavelength λ and frequency f. This is *quantum hypothesis*. So a given mode may have energies, $0E_0$, $1E_0$, $2E_0$, $3E_0$,...,nE_0, that is, $E = nhf$ and $n = 0, 1, 2, 3,...$. He also assumed that $E_0 = hf$. Planck then calculated both the probability of exciting a particular energy of a mode and the total energy of that mode. From these expressions, he was able to derive the relation

$$I_f = (2h)/(c^2)[(f^3)/(e^{hf/kT} - 1)]. \tag{22.19}$$

This is Planck's law. Here, h is Plank's constant and c the speed of light in a vacuum. Equation 22.19 is expressed in terms of the frequency of a mode and is the usual way Planck's law is written. It can be expressed in terms of λ, then Equation 22.19 becomes

$$I_\lambda = \frac{(2c)h}{\lambda^3}\left(\frac{1}{e^{hc/\lambda kT} - 1}\right). \tag{22.20}$$

Equation 22.20 is a good fit to the data of Figure 22.7 for all wavelengths. Although not evident without detailed mathematical analysis, Planck's hypothesis had the effect of putting a "cut-off " on the average energy per frequency mode emitted, so that it does not obey the classical assumptions of continuous, unbounded energy. The spectrum is bounded as $f \to \infty$.

EXAMPLE 22.9

A simple, ingenious way to make a blackbody ($e = 1$) emitter is to drill a small hole into the side of a chunk of metal. Standing waves will be set up in the drilled hole, so it will behave as a cavity. Suppose a piece of tungsten, with a drilled hole, is heated to a temperature of 2000 K. Radiations from both the tungsten surface and hole are measured. The emissivity of tungsten at 2000 K is $e = 0.28$. Determine (a) the wavelength at which the intensity R is a maximum for both the tungsten surface and the hole, (b) the value of R, at λ_{max}, for the tungsten and (c) for the hole, and (d) I_λ at λ_{max} for the hole.

Solution

a. To determine λ_{max}, the Wien displacement law gives

$$\lambda_{max} = \frac{2.898 \times 10^{-3}\,\mathrm{mK}}{T} = \frac{2.898 \times 10^{-3}\,\mathrm{mK}}{2000\,\mathrm{K}} = 1.45 \times 10^{-6}\,\mathrm{m}.$$

b.

$$R_{Tung} = \sigma e T_0^4 = \left(5.67 \times 10^{-8}\,\frac{\mathrm{W}}{\mathrm{m^2K^4}}\right)(0.28)(2000\,\mathrm{K})^4 = 2.5 \times 10^5\,\frac{\mathrm{W}}{\mathrm{m^2}}.$$

c. Here, e = 1

$$R_{Tung} = \left(5.67 \times 10^{-8} \frac{W}{m^2K^4}\right)(1.0)(2000\,K)^4 = 9.1 \times 10^5 \frac{W}{m^2}.$$

d.

$$I_f = (2hc)/\lambda^3[1/(e^{hc/(\lambda kT)} - 1)] = 9.07 \times 10^{-10}\,W/m^2.$$

Note that R_{hole} in part (c) is the intensity emitted for all wavelengths, whereas I_λ in part (d) is the intensity emitted only from an infinitesimal range around λ_{max}.

Originally, Planck's quantum hypothesis was considered by many physicists as an interesting "trick," but of no fundamental importance. All physical processes were considered to be continuous. Nevertheless, with the derivation of the Planck radiation law, the fourth revolution, that is, the quantum theory had begun.

22.10 WAVE–PARTICLE DUALITY

An insightful experiment in the early twentieth century showed that particles (electrons) can also behave as waves. An essential feature of waves is that they show interference effects. In Young's double-slit experiment, coherent light is incident on two transparent, closely spaced narrow slits, cut in an opaque barrier. The transmitted light creates a series of bright and dark fringes on a screen placed "downstream" from the slits.

Correspondingly, if a beam of electrons is directed onto two slits cut in a barrier, the transmitted electrons create a series of bright and dark fringes on a specially constructed screen (Figure 22.9).

The screen is similar to a television picture tube so that it glows when struck by electrons. The fringe pattern implies that the electrons are exhibiting interference effects that are characteristic of waves.

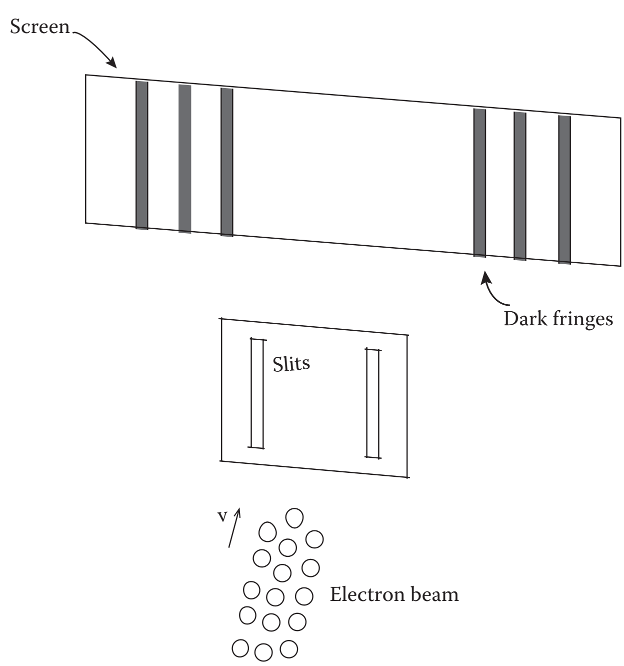


FIGURE 22.9 The double-slit experiment with a beam of electrons instead of light.

So, perceiving the electron as a small, discrete particle of matter does not account for or explain its apparent wave-like behavior in some experiments. Apparently, the electron can exhibit both particle-like and wave-like characteristics. That is, it exhibits a dual nature.

If particles (electrons) can behave as waves, can waves behave as particles? Indeed, prior to the above electron beam, double-slit, interference experiment, it was shown that a beam of light directed against a metal target can cause the ejection of electrons from the metal surface. This is called the *photoelectric effect* (see Figure 22.10).

The metal is placed in an evacuated tube with a voltage source between it and a positive collector electrode. If the photons have sufficient energy, electrons will be emitted from the metal, be attracted to the collector electrode, and thus constitute a current in the circuit. The emitted electrons, called *photoelectrons*, are due to their interaction with the photons of the incoming light wave. The interaction is similar to a particle–particle collision. Although EMWs appear continuous, the photoelectric effect indicates that they can be perceived as a beam composed of discrete photons.

If EMWs can behave as particles, they must possess or carry energy and linear momentum. Analysis of the photoelectric effect and the Compton effect, where x-ray photons collide with an electron at rest, resulting in a recoil electron and a scattered x-ray photon, yield expressions for these quantities. They are summarized in Table 22.2.

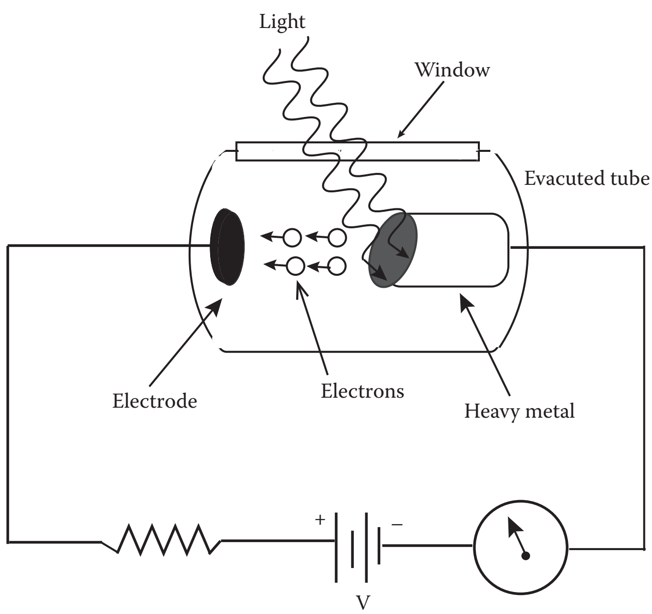


FIGURE 22.10 Light impinging on a metal target can cause it to eject electrons.

TABLE 22.2
Wave-Particle Duality

Quantity	Particle	Photon
Energy	$\frac{1}{2}mv^2 + PE$	$E = hf$
Linear momentum	$p = mv$	$p = \frac{h}{\lambda}$
Wavelength	$\lambda = \frac{h}{p} = \frac{h}{mv}$	$\lambda = \frac{h}{p}$

Since photons have zero mass when at rest, that is, when they are absorbed, the equations for the linear momentum and the kinetic energy of a photon cannot contain a mass term. Also, the expression for the wavelength of a particle, $\lambda = h/(mv)$, is called the *de Broglie wavelength* after Louis de Broglie, who proposed it in 1923.

EXAMPLE 22.10

Determine the de Broglie wavelength for (a) a proton ($m_p = 1.67 \times 10^{-27}$ kg) moving with a speed 4.00×10^6 m/s and (b) a baseball (mass = 0.150 kg) moving at a speed of 85.0 miles/h = 38.0 m/s.

Solution

a.

$$\lambda = \frac{h}{p} = \frac{h}{m_p v} = \frac{6.63 \times 10^{-34}\,\text{Js}}{(1.67 \times 10^{-27}\,\text{kg})(4.00 \times 10^{6}\,\text{m/s})} = 9.93 \times 10^{-14}\,\text{m} = 9.93 \times 10^{-12}\,\text{cm}.$$

For comparison, the size of an atom is of order 10^{-8} cm and an atomic nucleus of order 10^{-13} cm. The proton above has a wavelength approximately 10 times larger than a nucleus.

b.

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34}\ \text{Js}}{(0.150\,\text{kg})(38.0\,\text{m/s})} = 1.16 \times 10^{-34}\,\text{m}.$$

This wavelength is so small that the diffraction pattern of a baseball passing through, say, a 1.0-m doorway (the "slit") is not observable. (Recall that $m\lambda = a \sin \theta_D$.)

The ability to calculate the de Broglie wavelength (λ_D) is novel, but what does it mean? For photons treated as particles, which are the dual of EMWs, the wavelength is associated with oscillating electric and magnetic fields. For material particles, treated as waves, what is oscillating?

The modern interpretation is that the wave is a *probability wave*. To see this, consider the electron double-slit experiment depicted in Figure 22.9. Recall that the electrons pass through the slits, hit the screen, and cause it to glow. When this process is initially started, no interference pattern is discernable on the screen. As time elapses, and the pulses on the screen are summed, the typical pattern of Young's bright and dark fringes emerges. The bright fringes occur where many electrons hit the screen, the dark fringes indicate few, if any collisions. Thus, the fringes represent the probability that the electron will be located at that spatial point. So the magnitude of the particle wave (actually the square of the magnitude) at a given point in space indicates the probability that the particle is at that point. Then what is oscillating? probability!

For the case of EMWs, the intensity of the fringes is proportional either to the square of the wave's electric field or to the square of its magnetic field. In a similar fashion for particle waves, the probability of the particles spatial location is proportional to the square of a quantity called its *wave function*, which is usually represented by the symbol Ψ (psi). Thus, material particle waves are related to probability waves.

22.11 QUANTUM MECHANICS

Quantum mechanics represents the evolution of physics theory as applied to microscopic matter, electrons, atom, molecules, and so on, and EMWs. Since matter possesses wave properties and their de Broglie waves are related to the probability of the matters spatial location, a method is needed to calculate this probability. The method for this calculation is central to quantum mechanics.

In the mid-1920s, the Austrian physicist Erwin Schrodinger (1887–1961) constructed a differential equation called the *Schrodinger equation* whose solution yields the wave function Ψ. The magnitude of the square of Ψ gives the probability. Schrödinger's differential equation is rather abstract and is beyond the level of this chapter, as are the manipulations with Ψ. Once Ψ is known, it can be used to calculate average values of the particles energy, linear momentum, position, velocity, and so on. Note that the simple formula for the de Broglie wavelength ($\lambda = h/p$) is applicable to free particles. If a particle also has a potential energy, for example, an electron in a hydrogen atom, then Ψ is the appropriate quantity for probability calculations.

Also, in the mid-1920s, the German physicist Werner Heisenberg (1901–1976) simultaneously and independently developed a formalism for determining the average values of the dynamical variables of matter. His theory was originally called *matrix mechanics* because it was cast in terms of matrix equations, not differential equations.

It has since been shown that these two methods are equivalent and yield the same results. Separately, each is an approach to quantum mechanics.

The Schrodinger equation for calculating Ψ is now widely used and the results from it form the basis of our current understanding of electrons, nuclei, atoms, molecules, and microscopic physics in general. For atoms, it supersedes the Bohr theory.

22.12 HEISENBERG UNCERTAINTY PRINCIPLE

A particle beam (electrons) passing through two narrow slits (Young's double-slit experiment) exhibits an interference pattern similar to that for light waves (Figure 21.4). Similarly, if the beam passes through a single narrow slit, it exhibits a diffraction pattern (Figure 22.11) also similar to that for light waves as shown in Figure 21.7. Such a diffraction pattern can be used to discern an important property of matter and matter waves.

Figure 22.11 shows an electron beam directed in x direction. The beam passes through a narrow slit of width "a" and forms a diffraction pattern on a screen. The central bright fringe is wider than the slit width a. The fringes represent the probability of the spatial location of the electrons and thus imply that the electrons, although directed along x, have a y component of linear momentum. This component is not known exactly, that is, it has an uncertainty in its value, which is labeled Δp_y. Also, the electron can pass through any region of the slit, so it has a spatial uncertainty labeled Δy

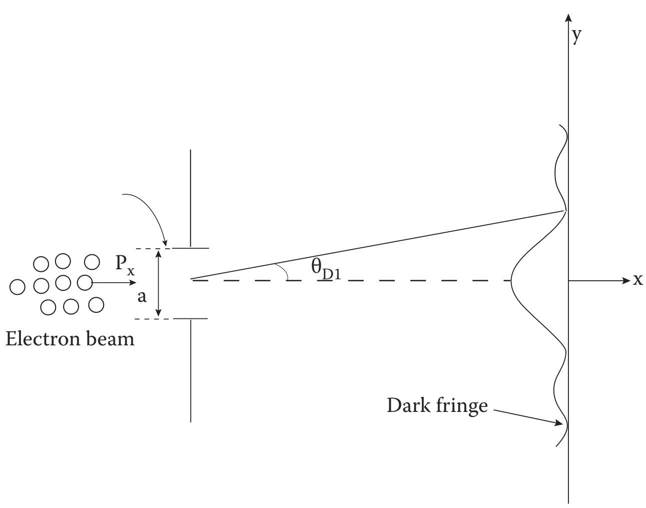


FIGURE 22.11 A single-slit diffraction pattern of electrons.

and here $\Delta y = a$. Consider the diffraction pattern in Figure 22.11. The condition to locate the dark fringes is, from Equation 21.19,

$$m\lambda = a \sin \theta_D. \tag{22.21}$$

Now $(\Delta y) \sim a$, so for $m = 1$,

$$\sin \theta_{D1} = \frac{\lambda}{a} = \left(\frac{1}{a}\right)\left(\frac{h}{p_x}\right) = \left(\frac{1}{\Delta y}\right)\left(\frac{h}{p_x}\right). \tag{22.22}$$

For small angles (<15°), $\sin \theta \sim \tan \theta$. Figure 22.11 gives

$$\sin \theta_{D1} \cong \tan \theta_{D1} = \left(\frac{\Delta p_y}{p_x}\right). \tag{22.23}$$

Equating the right sides of Equations 22.22 and 22.23 gives

$$\left(\frac{1}{\Delta y}\right)\left(\frac{h}{p_x}\right) \cong \frac{\Delta p_y}{p_x}.$$

Rearranging the above equation yields

$$(\Delta p_y)(\Delta y) \approx h. \tag{22.24}$$

A more thorough analysis of Equation 22.24 replaces the "approximate" symbol with an inequality. So, Equation 22.24 is replaced by

$$(\Delta p_y)(\Delta y) \geq \frac{h}{2\pi}. \tag{22.25}$$

Equation 22.25 is called the *Heisenberg uncertainty principle*. Equation 22.25 is a general principle even though it was derived for the special case of single-slit particle beam diffraction. It is general because particle waves are related to probabilities. So, when particles are described by a superposition of probability waves (a wave packet), their spatial position cannot be exactly specified, that is, $(\Delta y \neq 0)$ in general.

Correspondingly, if their spatial position is uncertain, so is their time rate of change of position (velocity) uncertain. Thus, $\Delta p_y \neq 0$ in general. Remember that Δy and Δp_y are "uncertainties." They represent by how much we are unsure of the exact values of position y and linear momentum p_y.

For motion in three dimensions, there are uncertainties in each of the three spatial coordinates and their associated linear momenta. In rectangular coordinates, this implies three relations of the form:

$$(\Delta x)(\Delta p_x) \geq \frac{h}{2\pi},$$

$$(\Delta y)(\Delta p_y) \geq \frac{h}{2\pi}, \tag{22.26}$$

$$(\Delta z)(\Delta p_z) \geq \frac{h}{2\pi}.$$

The units of the product, (coordinate) × (linear momentum), are (kg m²/s). This product is called *action*. The Heisenberg uncertainty principle applies to quantities whose product is "action." This leads to the additional equations:

$$(\Delta E)(\Delta t) \geq \frac{h}{2\pi} \tag{22.27}$$

and

$$(\Delta L)(\Delta \theta) \geq \frac{h}{2\pi}. \tag{22.28}$$

In Equation 22.27, (ΔE) is the uncertainty in the measured energy which was done during the time interval (Δt). In Equation 22.28, (ΔL) is the uncertainty in the magnitude of the angular momentum and (Δθ) is the uncertainty in the direction of the angular momentum vector.

EXAMPLE 22.11

The diameter and x component of the linear momentum of a B*B, shot from a B*B gun, are measured simultaneously. The diameter is found to be (4.50 ± 0.01) mm. Determine the value of the minimum uncertainty in its linear momentum.

Solution

Use

$$(\Delta x)(\Delta p_x) \geq \frac{h}{2\pi}.$$

Here, $(\Delta x) = 0.01\ \text{mm} = 1.0 \times 10^{-5}$ m and minimum (Δp_x) implies the equality sign applies. Thus,

$$(\Delta p_x) = \frac{h}{2\pi(\Delta x)} = \frac{(6.63 \times 10^{-34}\ \text{Js})}{2\pi(1.0 \times 10^{-5}\ \text{m})} = 1.1 \times 10^{-29}\ \frac{\text{kg m}}{\text{s}}.$$

This value is extremely small and shows that the limitations placed on the simultaneous determination of its position and linear momentum, by Equations 22.26, for such a macroscopic object, are inconsequential.

22.13 ATOMIC NUCLEUS, RADIOACTIVITY

Atoms consist of electrons, in orbit, revolving about a central nucleus. The allowed electron orbits are determined by the quantum mechanical rules on the quantum numbers n, l, m_ℓ, and m_s, as discussed in Section 22.5.

Experiments regarding the nature of the atomic nucleus began after the rather serendipitous discovery of radioactivity by the French chemist Henri Becquerel (1852–1908) in 1896.

The nucleus of an atom is made up of two main particles: *protons* and *neutrons* that are collectively called *nucleons*. The neutron, discovered by the English physicist James Chadwick (1891–1974), possess no electrical charge and has a mass $m_n = 1.674 \times 10^{-27}$ kg. It is 1837 times more massive than the electron. The proton has a positive charge $q_p = 1.60 \times 10^{-19}$ C that is equal in magnitude to the negative charge on an electron and a mass $m_p = 1.672 \times 10^{-27}$ kg that is approximately equal to the neutron mass. In electrically neutral atoms, the number of protons must be equal to the number of electrons. The symbol Z (atomic number) typically represents the number of protons in an atom. The symbol N represents the number of neutrons. If A represents the *mass number* or *nucleon number*, then

$$A = Z + N. \tag{22.29}$$

A shorthand notation has been devised that displays A and Z explicitly. To illustrate, the element iron with chemical symbol Fe has 26 protons and 56 nucleons. The shorthand notation would display this as $\left({}^{56}_{26}\text{Fe}\right)$. The generalized form of this notation is

$${}^{A}_{Z}X, \tag{22.30}$$

where X should be replaced by the chemical symbol of the element. To determine the number of neutrons from Equation 22.30, just form $N = A - Z$.

For iron, $N = 56 - 26 = 30$ neutrons. The notation of Equation 22.30 can be applied separately to protons, neutrons, and also electrons. Since a proton is the nucleus of an atomic hydrogen atom (chemical symbol = H), it may be denoted $({}^{1}_{1}\text{H})$. A neutron is displayed as $({}^{1}_{0}\text{n})$ and an electron as $({}^{0}_{-1}\text{e})$. For the electron, $A = 0$, since it contains no protons or neutrons, and $Z = -1$.

Nuclei, of a given chemical element, must all possess the same number of protons but may contain a different number of neutrons. These are called *isotopes*. For example, hydrogen has three known isotopes:

$$\text{(Hydrogen)},\ N = 0,\ Z = 1,\ A = 1,\ \text{symbol } ({}^{1}_{1}\text{H}).$$

$$\text{(Deuterium)},\ N = 1,\ Z = 1,\ A = 2,\ \text{symbol } ({}^{2}_{1}\text{H}).$$

$$\text{(Tritium)},\ N = 2,\ Z = 1,\ A = 3,\ \text{symbol } ({}^{3}_{1}\text{H}).$$

A second example is carbon that has two stable isotopes where $N = 6$ or $N = 7$. Since $Z = 6$, $A = 12$ or 13, the isotopes are $({}^{12}_{6}\text{C})$ (98.90% occurrence in nature) and $({}^{13}_{6}\text{C})$ (1.10%). Periodic tables show the mass number for an element that is the weighted average of its isotopes natural abundance. Protons and neutrons are packed together to make the shape of a nucleus approximately spherical (see Figure 22.12).

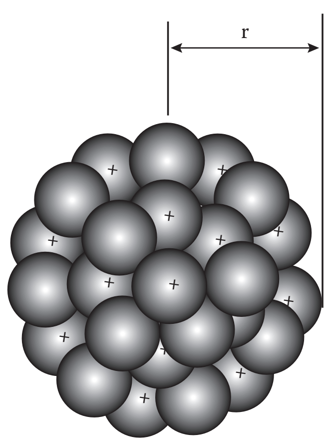


FIGURE 22.12 Atomic nuclei are approximately spherical in shape.

Experiments indicate that the radius of the nucleus depends on A and is approximately given by the equation

$$r \approx (1.2 \times 10^{-15}\,\mathrm{m})(A)^{1/3}. \tag{22.31}$$

EXAMPLE 22.12

Determine the (a) radius, (b) volume, and (c) density of the iron-56 ($^{56}_{26}Fe$) nucleus.

Solution

a. Using Equation 22.31

$$r \approx (1.2 \times 10^{-15}\,\mathrm{m})A^{1/3} = (1.2 \times 10^{-15}\,\mathrm{m})(56)^{1/3} = 4.6 \times 10^{-15}\,\mathrm{m}.$$

b.

$$V = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi(4.6 \times 10^{-15}\,\mathrm{m})^3 = 4.1 \times 10^{-43}\,\mathrm{m}^3.$$

c.

$$\rho = \frac{m}{V} = \frac{(26m_p + 30m_n)}{V} \approx \frac{56m_p}{V} = \frac{56(1.67 \times 10^{-27}\,\mathrm{kg})}{4.1 \times 10^{-43}\,\mathrm{m}^3} = 2.3 \times 10^{17}\,\mathrm{kg}.$$

Actually, the nuclei of all elements have approximately equal densities. The density of a lead nucleus is no greater than the density of a helium nucleus. This is because nuclear densities do not depend on the mass number A. To see this, use

$$\rho \equiv \frac{m}{V} \approx \frac{Am_p}{(4/3)\pi r^3} = \frac{Am_p}{(4/3)\pi(1.2 \times 10^{-15}\,\mathrm{m})^3 A}.$$

The mass number A appears in both numerator and denominator and divides out. So,

$$\rho \approx \frac{m_p}{(4/3)\pi(1.2 \times 10^{-15}\mathrm{m})^3} \cong 2.3 \times 10^{17}\,\frac{\mathrm{kg}}{\mathrm{m}^3},$$

which holds for all nuclei.

22.13.1 Stability

How is a nucleus held together since its neutrons are uncharged, but its positively charged protons should repel each other according to Coulomb's law, Equation 13.4. The nucleus should be unstable. For the closely packed nucleons to remain "bunched," that is, held together, there must exist some other attractive force to balance or offset the Coulomb electrostatic force. This force, called the *strong nuclear force*, is one of three fundamental forces found in nature.

Although not totally understood, some properties of this strong nuclear force are known. These properties are as follows:

1. It is almost independent of charge. For a given nucleon separation, the attraction between two protons, two neutrons, or between a neutron and a proton are nearly equal.
2. It is a very short range force. The attraction between two nucleons is very strong when their separation is of the order of 10^{-15} m and drops off to zero at larger separations. Because of this short range, only nearest neighbor nucleons are attracted to each other. In contrast, the electrostatic force varies as the reciprocal of the separation squared, and so is a much longer range force. So one proton repels all the other protons in the nucleus whereas it, or a neutron, attracts only its nearest neighbors.

These properties have an effect on nuclear stability. Referring to the periodic table, as the number of nuclear protons, and hence the repulsive force increases, the number of neutrons must increase to add to the attraction, but not the repulsion between nucleons. For stability, it is found that generally as Z increases, the number of neutrons exceeds the number of protons, that is, $N > Z$. Upon increasing Z, a point is reached where the balance between repulsive and attractive forces cannot be maintained by increasing N. The nucleus becomes unstable. This occurs for $Z > 83$.

The largest stable nucleus has 83 protons and is bismuth ($^{209}_{83}$Bi) with $(209 - 83) = 126$ neutrons. All nuclei with $Z > 83$ can occur in nature, but are unstable and either rearrange their internal structure over time or spontaneously break apart. That is, they are *radioactive*.

Stable nuclei are held together by the strong nuclear force. To separate such nuclei into their constituent nucleons requires energy. The energy holding the nucleons packed together is called the *binding energy*. The more stable a given nucleus, the greater is its binding energy. To determine the binding energy, it is experimentally observed that the mass of the intact nucleus is less than the sum of the masses of its constituent nucleons. This mass difference is called the mass defect and is designated Δm, where

$$\Delta m = [(Zmp + Nmn) - MA].$$

The theory of relativity states that rest mass and energy are equivalent. So, once Δm is determined, the binding energy can be calculated from the famous equation $E = mc^2$.

Applied here,

$$\text{Binding energy} = BE = (\Delta m)c^2. \tag{22.32}$$

EXAMPLE 22.13

Determine the binding energy of the Beryllium nucleus of "intact" mass, $M_A = 1.4964 \times 10^{-26}$ kg.

SOLUTION

$$\text{Bi has } Z = 4 \text{ and } N = 5\text{, so } (\Delta m) = 4\ mp + 5\ mn - MA$$

or

$$(\Delta m) = \{[4(1.673) + 5(1.675)] - 14.964\} \times 10^{-27}\,\text{kg} = (0.103) \times 10^{-27}\,\text{kg},$$

so,

$$BE = (\Delta m)c^2 = (0.103 \times 10^{-27}\text{ kg})(3.00 \times 10^8)^2 = 9.27 \times 10^{-12}\text{ J}$$

or

$$BE = (9.27 \times 10^{-12}\text{ J})\left(\frac{1.00\text{ eV}}{1.60 \times 10^{-19}\text{ J}}\right) = 5.79 \times 10^7\text{ eV} = 57.9\text{ MeV}.$$

Note the contrast between the large BE of nucleons, in millions of electron volts, and the ionization energy (a form of binding energy) of electrons, which is several electron volts.

22.13.2 Radioactivity

Unstable, or radioactive, nuclei rearrange their internal structure or break apart, that is decay, by emitting energy in the form of either particles or high-energy photons. Three kinds of particles/photons are possible from naturally occurring radioactive decay. They are *alpha* (α) particles, *beta* (β) particles, and *gamma* (γ) rays.

Alpha (α) particles are equivalent to the nuclei of doubly ionized helium (^4_2H). That is, they consist of two protons and two neutrons clumped together. They have a charge of Z = +2 and mass number A = 4. Because of their charge, αs are not very penetrating and travel only a few inches in air before being attracted and absorbed by other atoms/molecules.

Beta particles come from nuclei and have the same mass as electrons but can occur naturally with either negative or positive charge. Symbolically, (β^-) represents a β particle with the same mass and charge as an electron. (β^+) is a *positron* with the same mass, but whose charge is equal in magnitude but opposite in sign to the electron. So Z = ±1 for βs.

Gamma (γ) rays are photons (EMW), uncharged, and oscillating electric and magnetic fields. In terms of penetrating matter, αs are the least penetrating, βs more so, and γs very penetrating.

When a nucleus undergoes radioactive decay, the nucleus and its emitted particle/photon must obey the following five conservation laws: (a) mass/energy, (b) linear momentum, (c) angular momentum, (d) electric charge, and (e) nucleon number.

22.13.2.1 α Decay

Unstable nuclei that emit α particles are said to undergo α decay. An example of the α decay process is the disintegration of thorium into radium plus an α particle. Symbolically;

$$^{230}_{90}\mathrm{Th} \rightarrow {}^{226}_{88}\mathrm{Ra} + {}^{4}_{2}\mathrm{He}. \tag{22.33}$$

By this notation, thorium is called the *parent* nucleus and radium, the *daughter* nucleus. The charge number of the thorium nucleus (90) is different than that of radium (88), so α decay converts one element into another. This is called *transmutation*. Notice that both charge number (Z) and mass number or nucleon number (A) are conserved in Equation 23.33.

If the symbol P represents the parent nucleus and D, the daughter nucleus, the general equation for α decay is

$$^{A}_{Z}\mathrm{P} \rightarrow {}^{A-4}_{Z-2}\mathrm{D} + {}^{4}_{2}\mathrm{He}. \tag{22.34}$$

The radium nucleus in Equation 22.33 is itself radioactive and undergoes α decay into ($^{222}_{86}\mathrm{Em}$), and this nucleus α decays into ($^{218}_{84}\mathrm{Po}$).

A radioactive nucleus that emits an α particle also releases energy to conserve mass/energy. Some of the original mass of the parent nucleus is converted into the kinetic energy of the daughter nucleus plus the α particle. Therefore, the sum of the masses of the daughter plus α is less than the parent mass.

To quantitatively determine these energies, it is conventional to define a new mass unit called the *unified atomic mass unit* (u). By definition, (1 u) is equal to one-twelfth of the mass of carbon-12, ($^{12}_{6}\mathrm{C}$). That is

$$1\ \mathrm{u} \equiv \left(\frac{1}{12}\right)\mathrm{mass}\left({}^{12}_{6}\mathrm{C}\right). \tag{22.35}$$

Numerically, along with its energy equivalent $E = mc^2$,

$$1\ u = 1.66053886 \times 10^{-27}\ \text{kg} \rightarrow 931.5\ \text{MeV}. \tag{22.36}$$

In terms of (u), and their energy equivalent in MeV, proton, neutron, and electron values are

$$mp = 1.00727646688\ u \rightarrow 938.3\ \text{MeV},$$

$$mn = 1.00866491560\ u \rightarrow 939.6\ \text{MeV},$$

$$me = 0.00054857990945\ u \rightarrow 0.5110\ \text{MeV}.$$

EXAMPLE 22.14

Determine (a) the energy released during the α decay of thorium $^{230}_{90}\text{Th}(230.0331\,\text{u})$ (Values for u taken from *Handbook of Chemistry and Physics*, 61st Edition, CRC Press, Boca Raton, FL, 1980–1981, p. B-326.) to radium $^{226}_{88}\text{Ra}(226.0254\,\text{u})$ and $^{4}_{2}\text{He}(4.0026\,\text{u})$, and (b) the velocities and kinetic energies of the daughter atom (Ra) and the α particle immediately after the decay occurs.

Solution

a. Since the sum of the masses of the daughter atom plus the α particle is less than the parent mass, that mass difference is the amount of energy released by the parent. It is carried away as kinetic energy by the daughter plus α (except for a very small amount carried away as a γ ray and neglected in such calculations). Thus,

$$\Delta m = [230.0331\ u - (226.0254\ u + 4.0026\ u)] = 0.0051\ u.$$

To determine the energy, use Equation 22.36 for the conversion factor:

$$E = (0.0051\,u)\left(\frac{931.5\ \text{MeV}}{1.0\ u}\right) = 4.8\ \text{MeV}.$$

b. The energy released by the parent atom is carried away by the daughter plus α. Energy conservation gives

$$E = \left(\frac{1}{2}\right)m_D v_D^2 + \left(\frac{1}{2}\right)m_\alpha v_\alpha^2. \tag{22.37}$$

There are two unknown quantities in Equation 22.37, so a second independent equation is needed. It can be acquired from linear momentum conservation. The parent atom is initially at rest, so its linear momentum is zero.

$$[\text{system momentum}]_{\text{before}} = [\text{system momentum}]_{\text{after}}$$

or

$$0 = m_D v_D + m_\alpha v_\alpha, \tag{22.38}$$

then

$$v_D = \left(\frac{-m_\alpha}{m_D}\right)v_\alpha. \tag{22.39}$$

Insertion of Equation 22.39 into Equation 22.37, plus some algebraic manipulation gives

$$\left(1-\frac{m_\alpha}{m_D}\right)v_\alpha^2 = \frac{2E}{m_\alpha}. \tag{22.40}$$

To solve Equation 22.40 for v_α, in units of m/s, it is necessary to convert the masses from u to kg and the energy from MeV to joules. Thus,

$$m_\alpha = (4.0026\ \text{u})\left(\frac{1.6605\times10^{-27}\ \text{kg}}{1.0\ \text{u}}\right) = 6.646\times10^{-27}\ \text{kg},$$

$$m_D = 3.753\times10^{-25}\ \text{kg} = 375.3\times10^{-27}\ \text{kg},$$

and

$$E = (4.8\times10^{6}\ \text{eV})\left(\frac{1.60\times10^{-19}\ \text{J}}{1.0\ \text{eV}}\right) = 7.6\times10^{-13}\ \text{J}.$$

These values in Equation 22.40 give,

$$\left(1-\frac{6.646}{375.3}\right)v_\alpha^2 = \frac{2(7.6\times10^{-13}\ \text{J})}{6.646\times10^{-27}\ \text{kg}}.$$

Rearranging and solving yields

$$v_\alpha = 1.5\times10^{7}\ \text{m/s} = 0.05\ \text{c}.$$

Equation 22.39 gives

$$v_D = \left(\frac{-6.646}{375.3}\right)(1.5\times10^{7}\ \text{m/s}) = 2.7\times10^{5}\ \text{m/s}.$$

Also,

$$(KE)_\alpha = \frac{1}{2}m_\alpha v_\alpha^2 = \frac{1}{2}(6.646\times10^{-27}\ \text{kg})(1.5\times10^{7}\ \text{m/s})^2 = 7.5\times10^{-13}\ \text{J} \rightarrow 4.7\ \text{MeV}$$

and

$$(KE)_D = \frac{1}{2}m_D v_D^2 = \frac{1}{2}(375.3\times10^{-27}\ \text{kg})(2.7\times10^{5}\ \text{m/s})^2 = 1.4\times10^{-14}\ \text{J} \rightarrow 0.088\ \text{MeV}.$$

So,

$$(KE)_\alpha + (KE)_D = (4.7+0.088) \approx 4.8\ \text{MeV}.$$

The α particle receives a much larger share of the decay energy than does the daughter atom.

22.13.2.2 β Decay

Experimentally, β particles are found to have a mass equal to that of an electron. The charge carried by a(β^-) is negative and equal to that of an electron ($\beta^- \rightarrow -1.60 \times 10^{-19}$ C) while the charge carried by a (β^+), called a *positron*, is positive and equal in magnitude to an electron's charge, that is, ($\beta^+ \rightarrow +1.60 \times 10^{-19}$ C).

β particles are emitted by the nucleus. They are not atomic electrons (except in the rather rare cases of electron capture where an innermost atomic electron (K electron) is pulled into the nucleus, then reemitted). When the radioactive nucleus has a neutron-to-proton ratio (N/Z) that is too high for stability, it usually emits a (β^-). Unstable nuclei whose N/Z ratio is too low may emit a (β^+). In (β^-) emission, one of the nuclear neutrons decays into a proton plus a (β^-) so that charge is conserved. But the nuclear charge and neutron numbers are changed, so $Z \rightarrow (Z+1)$ and $N \rightarrow (N-1)$ and the N/Z ratio is lowered. *Note:* Neutrons themselves are unstable and a free neutron will decay into a proton and a (β^-) with an average lifetime of approximately 15 min.

In (β^+) emission, a nuclear proton is transformed into a neutron. As in α decay, both (β^-) and (β^+) decays result in transmutation. In (β^-), a neutron decays into a proton, so $Z \rightarrow (Z+1)$ whereas, in (β^+) decay, a proton decays into a neutron and $Z \rightarrow (Z-1)$.

It is found experimentally that most β particles generally have less kinetic energy than they should, to account for the energy released in the decay. It was suggested, in 1930 by Wolfgang Pauli, that another particle, called a *neutrino*, is emitted, along with β, and carries away some of the kinetic energy.

Thus, the general form for (β^-) decay, where the additional particle is called an *antineutrino* ($\bar{\upsilon}$), is

$$^{A}_{Z}\mathrm{P} \rightarrow {}^{A}_{Z+1}\mathrm{D} + {}^{0}_{-1}\mathrm{e} + \bar{\upsilon}. \tag{22.41}$$

For (β^+) decay, the additional particle is a neutrino (υ), so the general equation is

$$^{A}_{Z}\mathrm{P} \rightarrow {}^{A}_{Z-1}\mathrm{D} + {}^{0}_{1}\mathrm{e} + \upsilon. \tag{22.42}$$

The neutrino (υ) and its antiparticle ($\bar{\upsilon}$) have zero electrical charge and so are very difficult to detect. Experimental verification of the existence of the neutrino occurred in 1956. Later experiments indicate it has a mass that is just a small fraction of the mass of an electron. Much remains to be understood and verified about neutrinos/antineutrinos.

Beta and neutrino decay are associated with a force that in the past was called the weak nuclear force. The weak nuclear force and the electromagnetic force are thought to be separate manifestations of the more fundamental *electroweak force*. Current theory recognizes three fundamental forces in nature: the gravitational force, the electroweak force, and the strong nuclear force.

22.13.2.3 γ Decay

Nuclei can exist in various discrete energy states, in a fashion similar to the electrons in atoms. The lowest energy state is the ground state and states with higher energies are excited states. When a radioactive nucleus emits an α or a β particle, it is sometimes left in an excited energy state (generally denoted by an asterisk *). The excited nucleus can make a transition to a lower energy state by emitting a γ-ray photon. Gamma-ray photons come from the nucleus and are much higher in energy (typically 10 keV to 10 MeV) than photons emitted by electron energy transitions in atoms. No transmutation occurs, so γ decay does not yield a daughter nucleus.

The general relation for the γ-decay process is

$$^{A}_{Z}\mathrm{P}* \rightarrow {}^{A}_{Z}\mathrm{P} + \gamma. \tag{22.43}$$

EXAMPLE 22.15

An excited state of radioactive sodium ($^{24}_{11}Na^*$) emits a γ-ray photon of energy 0.423 MeV. Determine the photon's frequency and wavelength.

Solution

The excited nucleus makes a transition to a lower energy state by emitting 0.432 MeV of energy. Thus,

$$\Delta E = hf$$

or

$$f = \frac{\Delta E}{h} = \frac{(0.432\,\text{MeV})(1.60 \times 10^{-13}\,\text{J/1.0 MeV})}{6.63 \times 10^{-34}\,\text{J s}} = 1.02 \times 10^{20}\ \text{Hz},$$

then

$$\lambda = \frac{c}{f} = \frac{3.00 \times 10^{8}\,\text{m/s}}{1.02 \times 10^{20}\,\text{Hz}} = 2.94 \times 10^{-12}\,\text{m}.$$

Recall that nuclear dimensions are of the order of 10^{-15} m, so this wavelength corresponds to approximately 1000 nuclear diameters.

22.14 DECAY RATE, ACTIVITY

The decay of a given radioactive nucleus is spontaneous. That is, it occurs without any apparent external influence or cause. It is also random, that is, without a definite rule. But, these spontaneous, random decays occur with a constant probability. This probability is independent of the age of the nucleus or of any conditions external to it. So, in a collection of a number of the same species of radioactive nuclei, some will decay in short order while others will remain undecayed for a long span of time. The probability, per unit time, for any particular nucleus to decay is usually represented by the symbol λ and is called the *decay constant.*

To determine an equation for the rate of decay, consider a mass of radioactive material that contains N_0 undecayed nuclei at time $t = 0$. Since the decay probability is constant, the number of decays per second, at any time t, will depend on the number of undecayed nuclei N present. As nuclei decay, the *decrease* in N, that is, ΔN, will be proportional to the number of undecayed nuclei, thus

$$\frac{\Delta N}{\Delta t} = -\lambda N. \qquad (22.44)$$

The negative sign is to explicitly indicate that N is decreasing with time. Equation 22.44 can be solved (integrated) by use of calculus to give

$$N = N_0 e^{-\lambda t}. \qquad (22.45)$$

Here, N is the number of radioactive (undecayed) nuclei in the sample after a time t has elapsed. Equation 22.45 is called an exponential decay and is depicted in Figure 22.13.

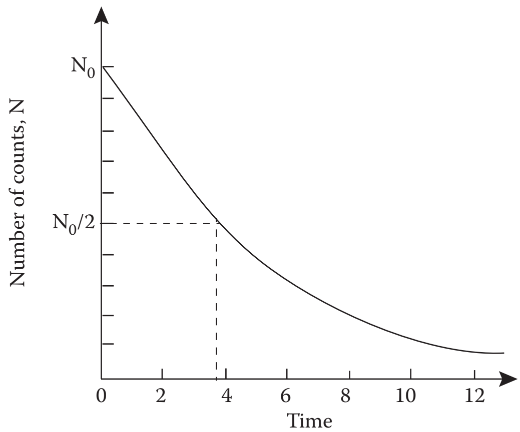


FIGURE 22.13 The activity of a radioactive material obeys an exponential decay curve.

The quantity (e) is the base of the natural logarithms and e = 2.718…. Some of the features of Figure 22.13 are as follows:

a. $N \to 0$ when $t \to \infty$.
b. $N = (N_0/2)$ at a particular time called the *half-life* of the radioactive material.

Item (a) implies that it takes infinitely long for all the radioactive nuclei in a given sample to decay. Item (b) allows a more meaningful characterization of the material, that is how much time must elapse so that the number of radioactive (undecayed) nuclei remaining is one-half of N_0.

To determine the half-life, let $N = (N_0/2)$ in Equation 22.45, then

$$\frac{N_0}{2} = N_0\, e^{-\lambda t_{1/2}}.$$

Dividing by N_0, taking the natural logarithm of both sides, and rearranging gives

$$T \equiv t_{1/2} = \frac{\ln 2}{\lambda} = \frac{0.693}{\lambda}. \tag{22.46}$$

The *activity* (A) of a radioactive material is a measure of the number of decays per unit time. This quantity is acquired by forming the time rate of change of Equation 22.45 and evaluating with calculus. That is,

$$A = \frac{\Delta N}{\Delta t} = \frac{\Delta}{\Delta t}\left(N_0\, e^{-\lambda t}\right) = N_0(-\lambda)e^{-\lambda t} \equiv A_0\, e^{-\lambda t}. \tag{22.47}$$

This equation may also be written as

$$A = -\lambda N. \tag{22.48}$$

EXAMPLE 22.16

A laboratory purchases a sample of radioactive gold ($^{196}_{79}Au$) containing (4.00×10^{16}) atoms. Its half-life is 6.18 days. Determine (a) the probability per second that any given gold nucleus will decay, (b) the activity at the end of one day and at the end of one half-life.

Solution

a. The probability per second is λ. The half-life is

$$T = (6.18\text{ days})\left(\frac{8.64 \times 10^4\text{ s}}{1.00\text{ days}}\right) = 5.34 \times 10^5\text{ s}.$$

So,

$$\lambda = \frac{\ln 2}{T} = \frac{0.693}{5.34 \times 10^5\text{ s}} = 1.30 \times 10^{-6}\text{ s}^{-1}.$$

b. The activity after 1.0 day = 24.0 h = 8.64×10^4 s is

$$A = -\lambda N_0 e^{-\lambda t} = (-1.30 \times 10^{-6}\text{ s}^{-1})(4.00 \times 10^{16})e^{-(1.30\times10^{-6}\text{ s}^{-1})(8.64\times10^4\text{ s})} = -4.65 \times 10^{10}\text{ counts/s}.$$

At the end of its half-life, $t = 5.34 \times 10^5$ s, so

$$A = -\lambda N = (-1.30 \times 10^{-6}\text{ s}^{-1})(4.00 \times 10^{16})e^{(-1.30\times10^{-6}\text{ s}^{-1})(5.34\times10^5\text{ s})} = -2.60 \times 10^{10}\text{ counts/s}.$$

The SI unit for activity is the *becquerel* (Bq) named after Henri Becquerel (1852–1908), a French chemist.

$$1.0\text{ Bq} \equiv 1.0\text{ disintegrations/s.} \quad (22.49)$$

Nuclear decays are measured, or counted, by particle detectors that sense the number of particles emitted. Thus, the word "counts" is associated with "disintegrations," so it is convenient to write,

$$1.0\text{ Bq} = 1.0\text{ counts/s} = 1.0\text{ cts/s.}$$

Historically the *curie*, named in honor of Marie (1867–1934) and Pierre (1859–1906) Curie, was the original unit of activity. By definition, the curie (Ci) is

$$1.0\text{Ci} \equiv 3.70 \times 10^{10}\text{ Bq.} \quad (22.50)$$

22.15 BIOLOGICAL CONSIDERATIONS

Ionizing radiation consists of photons, α particles, and β particles with sufficient energy to ionize atoms or molecules. That is, the radiation can, in a collision, knock electrons out of their atomic/molecular orbits. The energy needed to ionize atoms/molecules ranges from about 2.0 to 35.0 eV. The ionizing photons can be x-rays or γ rays and usually have energies in the 10 keV–8.0 MeV range, βs in the 10 keV–5.0 MeV range, or αs in the 10 keV–8.0 MeV range. Because of these high energies, a single α, β, or γ can ionize thousands of atoms/molecules. Because of this ionizing capability, radiation can alter the structure of molecules within living human cells, often resulting in the death of the cells. Thus, ionizing radiation can be harmful to human beings.

Correspondingly, if applied with proper precautions, ionizing radiation can be used for both medical diagnostics and therapy.

As an example of diagnostics, x-rays (photons) more readily pass through soft human tissue, such as muscles and organs, than through more densely packed cells such as bone or some cancerous tumors. X-rays can also be detected by photographic film. Thus, if a beam of x-rays is directed at a part of the human body, with a photographic plate behind it, the bone or tumor will absorb more of the beam and appear darker on the film than the soft tissue. Thus, a "negative" picture is formed, so a bone fracture or a tumor may be detected and analyzed.

As an example of therapy, Gamma Knife radiosurgery is a technique for treating problems of the brain, such as tumors and blood vessels. The technique is noninvasive, bloodless, and apparently painless and does not involve a physical knife. The procedure makes use of γ rays from a radioactive cobalt-60 source. The patient wears a γ ray absorbing metal helmet with small holes drilled through it. The holes focus the γ rays to a very small region where the tumor is located. Because of these needle beams of γ rays, most of the surrounding brain tissues experience minimal damage whereas the tumor, at the focus of these beams, receives a very large dose of radiation and is destroyed.

A second therapeutic example is the use of iodine-131 to kill cancerous tumors in the thyroid gland. Iodine-131, half-life 8.1 days, emits (0.606 MeV) β particles. When absorbed by the thyroid gland, the iodine concentrates in the cancerous tissue and the βs selectively kill the cancer cells. Because they grow and divide more rapidly than normal cells, cancer cells are more susceptible to death by radiation.

22.16 NUCLEAR REACTIONS

Recall that atomic nuclei with $Z \geq 83$ are unstable and will ultimately reach a stable state by radioactive decay. It is also possible to induce instability in a stable nucleus, even if it is relatively small with, say $A < 60$, by bombarding it with other nuclei, α particles, γ rays, neutrons, or protons. That is, a *nuclear reaction* occurs when the incident projectile particle causes a change to occur in the target nucleus. An example is the bombardment of carbon-13 atoms by high-energy protons. The reaction is

$$^{1}_{1}H + ^{13}_{6}C \rightarrow ^{14}_{7}N + \gamma.$$

Here, the proton remains in the nucleus, so that both nucleon and charge increase by (+1), and so induced transmutation occurs.

Note: Neutrons are a more expeditious choice of projectile. Since they are uncharged, they are not repelled by the electrostatic force of the nuclear protons. They can have much less kinetic energy than a proton or α, yet still penetrate the nucleus.

22.16.1 Nuclear Fission

If a heavy nucleus is bombarded with and absorbs a neutron, it may split into two smaller nuclei, each with a smaller mass than the target nucleus. This is known as *nuclear fission*. The example shown below is uranium-235 bombarded with neutrons.

$$^{1}_{0}n + ^{237}_{92}U \rightarrow ^{236}_{92}U* \rightarrow ^{141}_{56}Ba + ^{92}_{36}Kr + 3^{1}_{0}n. \quad (22.51)$$

A large amount of energy is released by the fission process. This energy was stored as electrostatic potential energy in the target nucleus and results in the kinetic energy of the fragments. Approximately 200 MeV per fission is released.

Nuclear reactors that can produce the steam used in the production of electricity or propulsion, and operate on the principle of controlled chain reaction nuclear fission. Notice that, in the specific reaction shown in Equation 22.51, three neutrons result from bombardment by a single neutron. Those three neutrons can keep the process going, a condition known as a *chain reaction*.

EXAMPLE 22.17

How many neutrons are released by the fission reaction

$$^1_0n + ^{235}_{92}U \rightarrow ^{133}_{51}Sb + ^{99}_{41}Nb + \text{neutrons?}$$

Solution

The projectile neutron is initially absorbed by the ($^{235}_{92}U$) atom, becomes unstable, and splits into antimony (Sb) and niobium (Nb) plus neutrons,

$$^1_0n + ^{235}_{92}U \rightarrow ^{236}_{92}U* \rightarrow ^{133}_{51}Sb + ^{99}_{41}Nb + x(^1_0n).$$

To conserve mass (nucleon) number A,

$$A = 236 = (133 + 99 + x).$$

So, $x = (236 - 232) = 4$ neutrons.

22.16.2 Nuclear Fusion

In the fission process, a "heavy" nucleus is split into two less massive nuclei. The process whereby two very "light," low-mass nuclei are forced to combine into a single, "heavier," more-massive nucleus is known as *nuclear fusion*. The combined mass of the "light" nuclei, before fusion, are greater than the mass of the single "heavier" nucleus that results from the fusion. This mass difference determines the amount of energy released by the reaction.

The three isotopes of hydrogen, namely hydrogen (^1_1H), deuterium (^2_1H), and tritium (^3_1H), are the light nuclei of interest in fusion physics. Deuterium and tritium have relatively low-binding energies per nucleon (approximately 1.1 and 2.8 MeV, respectively), which is favorable to the fusion process. Fusion is difficult to achieve practically since two nuclei must be brought close enough to each other so that the attractive strong nuclear force is greater than the repulsive electrostatic force. Then, the strong nuclear force can pull the two nuclei together, resulting in fusion. The amount of energy per nucleon released by fusion (approximately 3.3 MeV) is greater than that released per nucleon by fission (approximately 0.9 MeV). To date, no commercial fusion reactors, for the production of energy, have been constructed.

EXAMPLE 22.18

How much energy (in MeV) is released by the fusion of two deuterium nuclei (^2_1H, $m = 2.0141\,u$) when the process yields a tritium nucleus (^3_1H, $m = 3.0161\,u$) and a proton (^1_1H, $m = 1.0078\,u$)?

Solution

$$m_b = \text{masses before fusion} = 2(2.0141\ u) = 4.0282\ u,$$

$$m_{after} = 3.0161\ u + 1.0078\ u = 4.0239\ u,$$

$$\Delta m = m_b - m_a = (4.0282 - 4.0239)\ u = 0.0043\ u.$$

So,

$$E = \left((0.0043\,u)\left(\frac{931.5\ \text{MeV}}{1.00\ u}\right)\right) = 4.0\ \text{MeV}.$$

Both fission and fusion release nuclear energy, yet the processes seem to be the reverse of each other. In the fission process, a large nucleus splits into two smaller nuclei, whereas with fusion, two light nuclei "fuse" into a larger nucleus. How can they both liberate nuclear energy? The answer is related to the change in binding energy per nucleon, (BE/A), before and after the process.

Recall that the sum of the masses of the individual nucleons, before they are assembled into a nucleus, is greater than the same nucleons after being combined in the nucleus. This mass difference, if multiplied by the speed of light squared, gives the BE of the nucleus. This relation was expressed previously by Equation 22.32, that is,

$$BE = (\Delta m)c^2 = [(Zm_p + Nm_n) - M_A]c^2.$$

The BE is the energy holding the nucleus together and must be supplied to break it into its separate constituent nucleons. The (BE/A) has been calculated for most elements and is depicted in Figure 22.14.

Recognize that nuclei with larger a (BE/A) have released more mass energy than nuclei with a smaller (BE/A) upon formation. Correspondingly, in both fission and fusion, the original nucleus/nuclei have *less* (BE/A) than the nuclei that result from the process. Thus, in both cases, energy is released. As an example for fission, a projectile neutron strikes a uranium-235 nucleus which may split into barium-141 and krypton-92. Both Ba and Kr have more (BE/A) than uranium-235, so energy is released. For fusion, two deuterons may fuse to yield either tritium plus a proton or helium-3 plus a neutron. In either case, the fusion product has more (BE/A), so, again, energy is released.

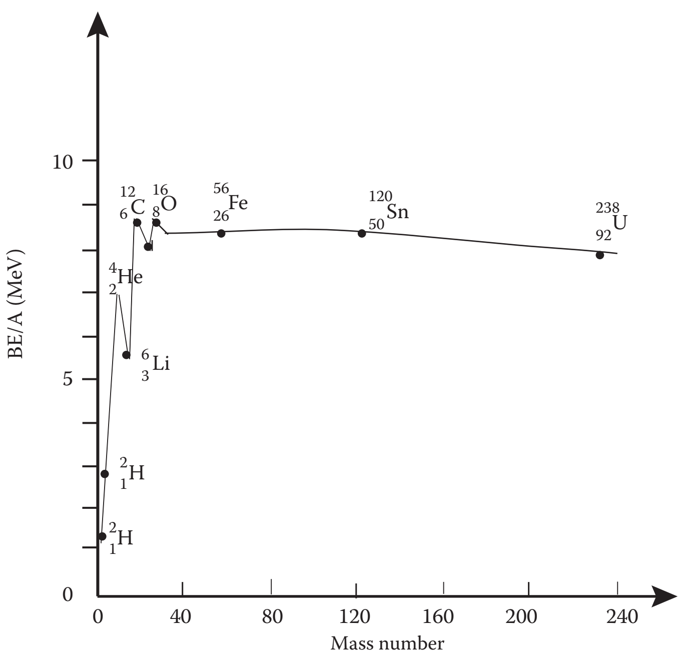


FIGURE 22.14 The binding energy per nucleon versus mass number for radioactive isotopes.

22.17 ELEMENTARY PARTICLES

Elementary particles are objects with no discernable internal structure. They are not composed or constructed from other particles.

One of the goals of physicists is to reveal and understand the fundamental building blocks, structures, processes, and rules of nature. To this end, many of the subcategories of physics, for example, classical mechanics, electromagnetism, optics, and thermodynamics, are reasonably well understood and are used successfully in applications and making predictions regarding the behavior of mechanical, electromagnetic, and heat phenomena.

Correspondingly, the existence, nature, properties, and rules of behavior of elementary particles are not completely understood. The study of elementary particles is at the forefront of our quest to determine the building blocks and understand both the microscopic and macroscopic nature of our universe.

By the early 1930s, experiments indicated that atoms (from the Greek word *atomos*, meaning indivisible) consisted of electrons, protons, and neutrons. These elementary particles, electrons, protons, and neutrons were considered the fundamental building blocks that constituted all matter. Their discovery had been made by investigating relatively low-energy (10 MeV or less) atomic/nuclear interactions such as atomic ionization and naturally occurring radioactivity.

With the invention and construction of charged particle accelerators, it became possible to bombard nuclei with higher energy projectiles and thus produce other "elementary" particles. These new particles are essentially unstable, with very short decay times between approximately 10^{-6} and 10^{-23} s. Most have masses greater than an electron and many exceed the mass of protons and neutrons. Just as there exists an antiparticle (the positron β^+) to the β particle (β^-), the new elementary particles have their associated antiparticles.

22.17.1 Classification

It is conventional to classify the elementary particles according to the type of force or forces by which they interact with other particles. The four fundamental interactions (forces) are shown in Table 22.3.

The second column in the table indicates relative strength. That means that within their range, as an example, the weak force has 10^{-6} or one one-millionth the strength of the strong force. Recall also that the strong force is responsible for the attraction of nucleons for each other and thus holds the nucleus together. The weak interaction is responsible for β decay, as an example, the conversion of a neutron into a proton. Both strong and weak interactions are very short ranged.

The grouping of particles by their interactions leads to three natural families or categories: (a) *photons*, (b) *leptons*, and (c) *hadrons*.

a. The photon category contains only one particle, the photon itself that interacts only with charged particles. Thus, the interaction is by the electromagnetic force.
b. Leptons are particles that do *not* interact with other particles by the strong force, but can interact via the three other forces. There exist six known leptons and their names and properties are shown in Table 22.4.
c. Hadrons are particles that interact via all four forces. But, at short distances ($\leq 10^{-15}$ m), the strong force dominates. Hadrons are also composite particles, made up of *quarks*, and therefore not really *elementary* particles. Among their members are protons, neutrons, and pions. It is conventional to subdivide hadrons into two subgroups: (a) baryons with half odd integer

TABLE 22.3
Four Fundamental Interactions

Interaction (Force)	Comparative Strength	Elementary Particles Affected	Range (m)
Strong	1	Quarks	10^{-15}
Electromagnetic	10^{-2}	Electrically charged	$1/(r^2) \to \infty$
Weak	10^{-6}	Quarks and leptons	10^{-17}
Gravitational	10^{-43}	All	$1/(r^2) \to \infty$

TABLE 22.4
Leptons

Name	Symbol	Charge (e)	Lifetime (s)	Rest Energy (MeV)
Electron	β^-	−1	Stable	0.511
Electron neutrino	υ_e	0	Stable	>0, but small
Muon	μ^-	−1	2.2×10^{-6}	105.7
Muon neutrino	υ_μ	0	Stable	<0.170
Tau	τ^-	−1	10^{-13}	1777
Tau neutrino	υ_τ	0	Stable	<15.5

spin (i.e., composite fermions) and (b) mesons with integer intrinsic spin (i.e., composite bosons). Note that elementary particles, such as electrons in the theory of the fine structure of atoms, can possess an intrinsic spin. Table 22.5 lists a few representative hadrons.

By inspection of the rest-energy column of Tables 22.4 and 22.5, it can be seen that leptons are the least massive, mesons are intermediate, and baryons are the most massive of the elementary particles.

22.17.2 Quarks

The constituent particles of hadrons are called quarks. They are the only known (assumed) carriers of fractional electron/positron charge. Because quarks and their antiparticles combine in groups of two to form mesons, only integer charge is observed in nature. Quarks combine in groups of three to form hadrons. Six quarks and their antiparticles are required to form the known hadrons. They are shown in Table 22.6.

To form a meson, a quark and an antiquark are required. For example, a π^+ meson is formed from $\pi^+ = u + \bar{d} = +(2/3)e + (1/3)e = +e$; that is, the charge of the π^+ is the same as a positron (β^+). To form a baryon, three quarks are required. For example, the proton is composed

$$p = u + u + d = \left(\frac{2}{3}\right)e + \left(\frac{2}{3}\right)e - \left(\frac{1}{3}\right)e = +e.$$

TABLE 22.5
Hadrons

Name	Symbol	Charge (e)	Lifetime (s)	Rest Energy (MeV)	Spin Quantum Number
			Baryons		
Proton	p	+1	Stable	938.3	1/2
Neutron	n	0	900	939.6	1/2
Sigma	Σ^+	+1	0.8×10^{-10}	1189	1/2
Omega	Ω^-	−1	0.8×10^{-10}	1672	1/2
			Mesons		
Pion	Π^+	+1	2.6×10^{-8}	139.6	0
Kaon	K^+	+1	1.2×10^{-8}	493.7	0
Eta	η^0	0	$<10^{-18}$	548.8	0

TABLE 22.6
Quarks

Name	Symbol	Charge (e)	Antiparticle	Antiparticle Charge (e)
Up	u	+(2/3)	$\bar{u}$	−(2/3)
Down	d	−(1/3)	$\bar{d}$	+(1/3)
Strange	s	−(1/3)	$\bar{s}$	+(1/3)
Charm	c	+(2/3)	$\bar{c}$	−(2/3)
Truth (top)	t	+(2/3)	$\bar{t}$	−(2/3)
Beauty (bottom)	b	−(1/3)	$\bar{b}$	+(1/3)

A neutron is

$$n = u + d + d = +\left(\frac{2}{3}\right)e - \left(\frac{1}{3}\right)e - \left(\frac{1}{3}\right)e = 0,$$

that is the neutron has zero electrical charge.

To date, no isolated, fractional charge quark has been experimentally detected or observed. Some scientists think they cannot be isolated because they are too tightly bound together by the strong nuclear force. Nevertheless, the quark model has been theoretically successful in accounting for the experimentally observed hadrons.

The currently accepted explanation of the fundamental forces, the strong nuclear force between quarks, and their formation of the hadrons is known as *the standard model*. This research is at the forefront of physics.

PROBLEMS

22.1 Determine which of the possible energy levels of aluminum (Al) its 13 electrons occupy in the ground state.

22.2 Estimate, using the Bohr model,
   a. The minimum energy required to remove an $n = 1$ electron from a platinum ($Z = 78$) target
   b. The wavelength of the K_α ($n = 2$ to $n = 1$) x-rays

22.3 Consider a proton ($m_p = 1.67 \times 10^{-27}$ kg) confined in an atomic nucleus of size $\Delta x \sim 1.00 \times 10^{-14}$ m. Assume the proton is nonrelativistic and determine
   a. The minimum magnitude of the protons linear momentum
   b. Its minimum speed
   c. Its minimum kinetic energy (in MeV) because of its confinement

22.4 A 2000-lb automobile ($m = 908$ kg) is traveling at 60.0 mi/h. The uncertainty in its velocity (due to speedometer error) is 2.00 mi/h. Determine the minimum uncertainty in the automobiles position as it is moving.

22.5 Determine the (a) radius and (b) density of the beryllium nucleus.

22.6 Determine the binding energies of (a) carbon-12 ($M_A = 1.9926 \times 10^{-26}$ kg) and carbon-13 ($M_A = 2.1592 \times 10^{-26}$ kg).

22.7 Determine the energy released by the α decay of uranium $^{238}_{92}U$(238.0508 u) to thorium $^{234}_{90}Th$(234.0436 u) and an α particle $^{4}_{2}He$ (4.0026 u). Express your answer in MeV.

22.8 The wavelength of the γ-ray photon, emitted by radioactive radium ($^{226}_{88}Ra^*$), is 6.67×10^{-12} m. Determine the energy (in MeV) of this photon.

22.9 In a certain house, it is found that 5.00×10^{10} radioactive radon atoms enter the basement through sewer drain pipes. The basement is then sealed so the atoms are trapped. The half-life of radon ($^{222}_{86}$Rn) is 3.83 days. Determine (a) the activity at the time the basement is sealed, (b) the activity 2 weeks later, and (c) the number of undecayed radon atoms at the end of 2 weeks.

22.10 How many neutrons are released by the fission reaction

$$^{1}_{0}\text{n} + ^{235}_{92}\text{U} \rightarrow ^{140}_{54}\text{Xe} + ^{94}_{38}\text{Sr} + \text{neutrons?}$$

22.11 A deuterium nucleus ($^{2}_{1}$H) collides with and is absorbed by a nitrogen nucleus ($^{14}_{7}$N). What is the element name, nucleon (mass) number A, and charge number Z of the resulting compound nucleus?

22.12 How much energy (in MeV) is released by the fusion of deuterium ($^{2}_{1}\text{H}, m = 2.014$ u) and tritium ($^{3}_{1}\text{H}, m = 3.016$ u)? The process yields a helium atom ($^{4}_{2}\text{He}, m = 4.003$ u) and a neutron ($^{1}_{0}\text{n}, m = 1.009$ u).

22.13 Balmer series: Determine the wavelength of the spectral line of atomic hydrogen resulting from a transition of its electron from (a) the n = 5 to the n = 2 levels, and (b) the n = 4 to the n = 3 levels. (c) Are these spectral lines in the visible region of the EM spectrum?

22.14 Calculate the energy of a visible range photon of wavelength $\lambda = 550.0$ nm. Express your answer in joules and MeV. (b) Calculate the frequency and wavelength of a hypothetical photon of energy 1.00 J.

22.15 Estimate the minimum energy that a beam electron must have to knock a ground-state electron (n = 1) completely out of a platinum target atom (Z = 78) orbit, in an x-ray tube. Use the Bohr model estimate.

22.16 A bottle of cold beer ($T_0 = 40.0°F = 277.7$ K) is taken from a refrigerator and placed on a table in a room of temperature $T_s = 72.0°F = 295.2$ K. The beer is left on the table and ultimately warms to room temperature. Assume it absorbs heat only by the radiation process. The bottle has a surface area of $A = 0.0427\ m^2$, a surface emissivity of e = 0.650, and the beer has a mass of 0.163 kg. Ignore the bottle's mass and assume a specific heat capacity for the beer of c = 4186 J/kg °C. Calculate the time in which the beer will reach room temperature.

22.17 Gold, at atmospheric pressure, melts at the temperature T = 1945°F = 1063°C = 1336 K. Assuming it is a blackbody radiator, calculate the wavelength of the maximum intensity radiated.

22.18 An oxygen molecule of mass 5.30×10^{-26} kg is confined within an alveoli sac in a persons lungs. The average diameter of these sacs is approximately 0.250 mm. Determine the minimum uncertainty in the molecules' (a) linear momentum, and (b) velocity, while in the sac.

22.19 Determine the (a) radius, (b) volume, and (c) mass density of the cadmium-112 ($^{112}_{48}$Cd) nucleus.

22.20 The atomic mass of cobalt ($^{59}_{27}$Co) is 58.933198 u. This is the mass of the atom, so it includes the masses of the (27) electrons. Calculate (a) the mass defect and (b) the binding energy for the cobalt nucleus. *Note:* The mass of atomic hydrogen is (1.0078 u), $m_n = 1.0087$ u, $m_p = 1.0073$ u, and $1.0\ \text{u} = 1.6605 \times 10^{-27}$ kg.

22.21 Radium ($^{226}_{88}$Ra, 226.02540 u) emits an α particle ($^{4}_{2}$He, 4.00260 u) when it decays to Radon ($^{222}_{86}$Rn, 222.01757 u). Determine the energy released by this α decay.

22.22 The half-life of cobalt-60 ($^{60}_{27}$Co, 58.93382 u) is 5.27 years. (a) Calculate its decay probability. (b) What will be the activity of a 1.00-g sample 1 week after purchase?

22.23 How many neutrons are released by the fission reaction

$$^{1}_{0}\text{n} + ^{235}_{92}\text{U} \rightarrow ^{236}_{92}\text{U}* \rightarrow ^{141}_{56}\text{Ba} + ^{92}_{36}\text{Kr} + x(^{1}_{0}\text{n})?$$

22.24 How much energy (in MeV) is released by the fusion of two deuterium nuclei ($^{2}_{1}H$, m = 2.0141 u) when the process yields helium-3 ($^{2}_{1}He$, m = 3.0160 u) and a neutron ($^{1}_{0}n$, m = 1.0087 u)? The reaction is $^{2}_{1}H + ^{2}_{1}H \rightarrow ^{3}_{2}He + ^{1}_{0}n$.

22.25 An antiproton ($\overline{p}$) has the same mass, but opposite charge, as a proton. Determine which quarks are required to compose an antiproton.

PHYSICS

ESSENTIAL PHYSICS

John Matolyak, Emeritus Professor, Indiana University of Pennsylvania, USA
Ajawad Haija, Indiana University of Pennsylvania, USA

Fluency with physics fundamentals and problem-solving has a collateral effect on students by enhancing their analytical reasoning skills. In a sense, physics is to intellectual pursuits what strength training is to sports.

Designed for a two-semester algebra-based course, **Essential Physics** provides a thorough understanding of the fundamentals of physics central to many fields. It omits material often found in much larger texts that cannot be covered in a year-long course and is not needed for non-physics majors. Instead, this text focuses on providing a solid understanding of basic physics and physical principles. While not delving into the more specialized areas of the field, the text covers the fundamentals of mechanics, electricity and magnetism, light, and modern physics.

This book is appropriate for a course in which the goals are to give the students a grasp of introductory physics and enhance their analytical problem-solving skills. Each topic includes worked examples. Math is introduced as necessary, with some applications in biology, chemistry, and safety science also provided. If exposure to more applications, special topics, and concepts is desired, this book can be used as a problem-solving supplement to a more inclusive text.

CRC Press
Taylor & Francis Group
an informa business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

K16686
ISBN-13: 978-1-4665-7521-9
90000
9 781466 575219