AS & A-Level Physics Lecture Notes

(Year 1)

BY

Yuhao Yang

Shanghai Easyday Education

August 31, 2020

Contents

Practical Issues

Last Update: August 31, 2020

Contacts

Email: colin-young@live.com

The latest update can be found via: https://github.com/yuhao-yang-cy/asphysics

About the Notes

This is a set of very concise lecture notes written for CIE AS-level Physics (syllabus code 9702). Presumably the target audience of the notes are students studying the relevant course.

These notes are supposed to be self-contained. I believe I have done my best to make the lecture notes reflect the spirit of the syllabus set by the Cambridge International Examination Board. Apart from the essential derivations and explanations, I also included a handful of worked examples and problem sets, so that you might get some rough idea about the styles of questions that you might encounter in the exams. If you are a student studying this course, I believe these notes could help you get well-prepared for the exams.

Despite the fact that the notes are supposed to serve the CIE candidates, the world of physical science is so rich and wonderful, and the lots of fascinating details of the nature and the universe produce insights and understandings of many things all around us. I would love to share and selectively include in the notes a small part that I know, although some of the materials are beyond the CIE requirements. I would apologize ahead as in some chapters I just could not help myself filling in materials that I find interesting. There are pages where footnotes take more space than the main body.

Many materials in the notes are borrowed from the textbooks listed on the next page and the know-it-all internet. I strongly recommend those readers who want to know more to take a close look at the list of references.

Throughout these notes, key concepts are marked red, key definitions and important formulas are boxed. The comments I would like to make on specific topics are followed by a left hand symbol (☞). Any non-examinable material that usually show up in the notes are pointed out explicitly. Anything labelled with a star sign is a warning sign that it is beyond the CIE syllabus.

I don't think anyone can learn physics without doing sufficient exercises, that is why I have offered a few worked examples. There are also two versions of the notes, one with an end-of-chapter question section for each chapter and one without. You can freely choose which version to satisfy your personal need, but I would recommend you to get your hands on as many problems as you can to test your understanding.

These lecture notes are not yet complete. I will follow them up with the development of the course.

Also very importantly, I am certain that there are countless typos in the notes. If you spot any errors, please let me know.

Literature

I borrow heavily from the following resources:

- Cambridge International AS and A Level Physics Coursebook, by *David Sang*, *Graham Jones*, *Richard Woodside* and *Gurinder Chadha*, Cambridge University Press
- International A Level Physics Revision Guide, by *Richard Woodside*, Hodder Education
- Longman Advanced Level Physics, by *Kwok Wai Loo*, Pearson Education South Asia
- Conceptual Physics (10^{th} Edition), by *Paul G. Hewitt*, Pearson International Education
- Fundamentals of Physics, by *Robert Resnick*, *David Halliday* and *Kenneth S. Krane*, John Wiley & Sons
- Past Papers of Cambridge Internation A-Level Physics Examinations
- HyperPhysics Website: http://hyperphysics.phy-astr.gsu.edu/hbase/index.html
- Wikipedia Website: https://en.wikipedia.org

Copyright

Chapter 1
Physical Quantities

1.1 units of measurement

any physical quantity contains a numerical value and its associated unit

a system of units of measurement used throughout the scientific world is the **SI units**[1]

1.1.1 SI base units

SI defines seven units of measure as a basic set, known as the **SI base units**

base quantity	base unit	symbol
mass	kilogram	kg
length	metre	m
time	second	s
electric current	ampere	A
temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd

1.1.2 derived units

the seven[2] SI base units are building blocks of the SI system

all other quantities are derived from the base units

[1]SI units, abbreviated from the French *Système Internationale d'Unités*, means the International System of Units. Those who are interested in the history and evolution of the SI can check out the Wikipedia article: https://en.wikipedia.org/wiki/International_System_of_Units

[2]Luminous intensity is beyond the scope of the AS-Level syllabus. You are only required to know the other six SI base quantities and their units.

Example 1.1 Give the SI base units of (a) speed, (b) acceleration, (c) force, (d) work done.

✍ speed $= \dfrac{\text{distance}}{\text{time}} \quad \Rightarrow \quad [v] = \dfrac{[s]}{[t]} = \dfrac{\text{m}}{\text{s}} = \text{m s}^{-1}$

acceleration $= \dfrac{\text{speed}}{\text{time}} \quad \Rightarrow \quad [a] = \dfrac{[v]}{[t]} = \dfrac{\text{m s}^{-1}}{\text{s}} = \text{m s}^{-2}$

force $=$ mass $\times$ acceleration $\quad \Rightarrow \quad [F] = [m][a] = \text{kg m s}^{-2}$

work $=$ force $\times$ distance $\quad \Rightarrow \quad [W] = [F][s] = \text{kg m s}^{-2} \times \text{m} = \text{kg m}^2\text{s}^{-2}$ □

1.1.3 metric prefixes

prefixes are used to indicate multiples and sub-multiples of original units

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

1.1.4 dimensional analysis

if an equation is correct, then the units on both sides must be the same

such an equation with consistent units is said to be **homogeneous**.

dimensional analysis is widely used as a rough guide to check for the correctness of equations

there are times when the dependence of one physical quantity on various other quantities cannot not be seen easily, but it might give us helpful hints by merely investigating their units

➢ there are *unit free*, or *dimensionless* quantities that do not have units

examples are real numbers (2, $\frac{4}{3}$, π, etc.), coefficient of friction (μ), refraction index (n), etc.

➢ a correct equation must be homogeneous, but the converse may not be true

possible problems include an incorrect coefficient, extra term, an incorrect sign, etc.

Example 1.2 A ball falls in vacuum, all its gravitational potential energy converts into kinetic energy. This is expressed by the equation: $mgh = \frac{1}{2}mv^2$. Show that this equation is homogeneous.

✍ LHS: $[mgh] = [m][g][h] = \text{kg} \times \text{ m s}^{-2} \times \text{m} = \text{kg m}^2\text{s}^{-2}$

RHS: $\left[\frac{1}{2}mv^2\right] = [m][v]^2 = \text{kg} \times (\text{ m s}^{-1})^2 = \text{kg m}^2\text{s}^{-2}$

so we see the equation $mgh = \frac{1}{2}mv^2$ is homogeneous □

Example 1.3 The speed of a wave travelling along an elastic string is determined by three things: the tension T in the string, the length L of the string, and the mass m of the string. Let's assume $v = T^a L^b m^c$, where a, b, c are some numerical constants. Find the values of a, b and c.

✍ RHS: $[T]^a[L]^b[m]^c = (\text{kg m s}^{-2})^a \text{m}^b \text{kg}^c = \text{kg}^{a+c}\text{m}^{a+b}\text{s}^{-2a}$

for the equation to be homogeneous, we must have:

$$\text{kg}^{a+c}\text{m}^{a+b}\text{s}^{-2a} = \text{m s}^{-1} \quad \Rightarrow \quad \begin{cases} \text{kg}: & a+c=0 \\ \text{m}: & a+b=1 \\ \text{s}: & -2a=-1 \end{cases} \quad \Rightarrow \quad \begin{cases} a=\frac{1}{2} \\ b=\frac{1}{2} \\ c=-\frac{1}{2} \end{cases}$$

so wave speed is given by: $v = T^{1/2}L^{1/2}m^{-1/2}$, or $v = \sqrt{\frac{TL}{m}}$

this happens to be the correct formula for the wave speed on a string □

1.2 scalars & vectors

physical quantities come in two types: scalars and vectors

a **scalar** quantity has magnitude only

a **vector** quantity has magnitude and direction

➢ a scalar can be described by a single number

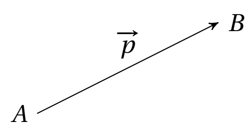
examples of scalars are time, distance, speed, mass, temperature, energy, density, volume, etc.

➢ a vector is usually represented by an arrow in a specific direction

a vector $\vec{p}$ pointing from A to B is shown

length of the arrow shows the magnitude of the vector

direction of the arrow gives the direction of the vector



examples of vectors are displacement, velocity, acceleration, force, field strength, etc.

➢ scalar algebra is just ordinary algebra

one can add and subtract scalar quantities in the same way as if they were ordinary numbers

for example, a set of objects with mass $m_1, m_2, \cdots, m_n$ have a total mass of $M = m_1 + m_2 + \cdots + m_n$

➢ vector algebra is more complicated, since we need keep track of the direction

1.2.1 multiplication of vectors

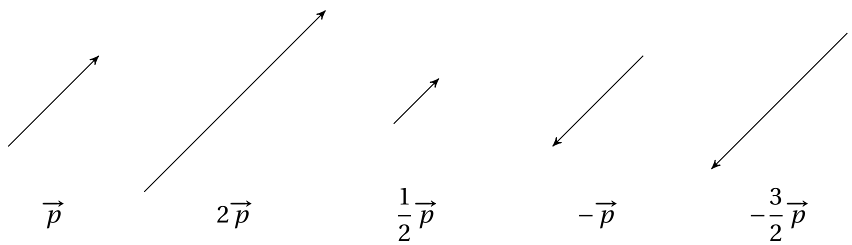
vectors can be multiplied by scalars [3]

when being multiplied by a scalar number, magnitude of the vector changes

if this number is positive, the vector becomes longer or shorter, but still points in same direction

if the number to be multiplied is negative, the operation reverses the vector's direction

Example 1.4 Given a vector $\overrightarrow{p}$, the graphical representations of $2\overrightarrow{p}$, $\frac{1}{2}\overrightarrow{p}$, $-\overrightarrow{p}$, $-\frac{3}{2}\overrightarrow{p}$ are:



1.2.2 addition of vectors

vectors can be added to form a **resultant** vector

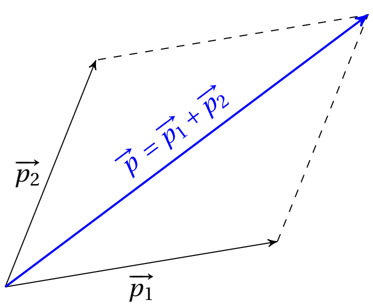
to deal with vector sums, we need take the directions of vectors into account

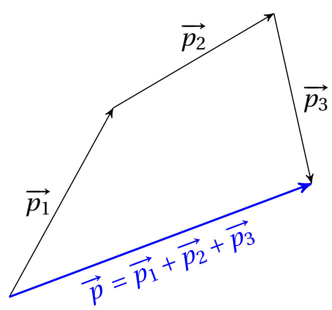
let's consider the sum of two vectors $\overrightarrow{p_1}$ and $\overrightarrow{p_2}$

resultant vector $\overrightarrow{p} = \overrightarrow{p_1} + \overrightarrow{p_2}$ lies on diagonal of the parallelogram subtended by $\overrightarrow{p_1}$ and $\overrightarrow{p_2}$

this is called the **parallelogram rule** for vector addition

if the resultant of several vectors $\overrightarrow{p} = \overrightarrow{p_1} + \overrightarrow{p_2} + \cdots \overrightarrow{p_n}$ is to be found, one can join these vectors head-to-tail, the resultant is given by the arrow connecting the tail of $\overrightarrow{p_1}$ to the head of $\overrightarrow{p_n}$

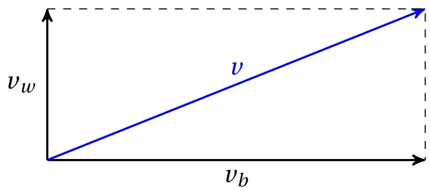




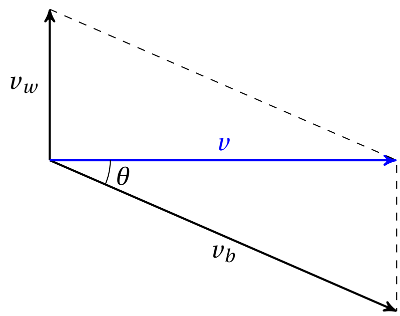
[3] It is also possible to multiply vectors with vectors, and there are basically two ways of doing vector multiplication: the *dot product* and the *cross product*. Both vector products are useful in physics, but we will not go into the details. You may learn more about vector multiplication in the A-Level mathematics course.

Example 1.5 A river flows from south to north with a speed of 2.0 m s^{-1} and the speed of a boat with respect to the water flow is 5.0 m s^{-1}. (a) Suppose the boat leaves the west bank heading due east, what is the resultant velocity of the boat? (b) If the boat is to reach the exact opposite bank across the river, what is the resultant velocity and in what direction should the boat be headed?

✍ vector diagrams for resultant velocity of the boat is illustrated below



(a) boat heading due south



(b) boat aiming at exact opposite bank

(a) magnitude of resultant velocity: $v = \sqrt{v_b^2 + v_w^2} = \sqrt{5.0^2 + 2.0^2} \approx 5.4$ m s^{-1}

in this case, the boat reaches opposite bank in shortest time but will drift downstream

(b) magnitude of resultant velocity: $v = \sqrt{v_b^2 + v_w^2} = \sqrt{5.0^2 - 2.0^2} \approx 4.6$ m s^{-1}

in this case, the boat reaches opposite bank in shortest distance

but the boat is headed slightly upstream: $\theta = \sin^{-1}\dfrac{v_w}{v_b} = \sin^{-1}\dfrac{2.0}{5.0} \approx 24°$ □

1.2.3 resolving vectors

one can also resolve a single vector into two (or more) components [4]

let's place a general 2D vector $\vec{p}$ in Cartesian coordinates

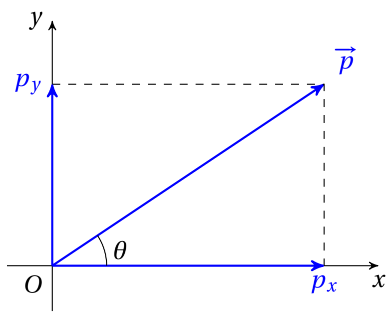
vector $\vec{p}$ can be split into two perpendicular components

– a horizontal component p_x

– a vertical component p_y

if $\vec{p}$ forms an angle θ to the x-axis, then:

$$p_x = p\cos\theta, \quad p_y = p\sin\theta$$
$$p = |\vec{p}| = \sqrt{p_x^2 + p_y^2}, \quad \tan\theta = \frac{p_y}{p_x}$$

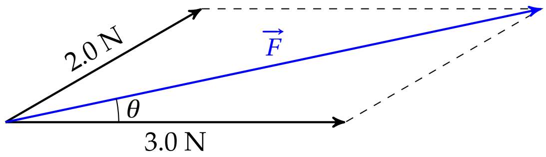


[4]This depends on the number of dimensions of space we are working with.

Example 1.6 A force of 3.0 N towards east and a force of 2.0 N towards 30° north of east act on an object. Find the magnitude and the direction of the resultant force.

✍ suppose an arrow of length 1 cm represents a force of 1 N

one can draw a *scale diagram* with a ruler and a protractor as shown



one can find length of the resultant vector is about 4.8 cm

also it forms an angle of about 12° to the 3.0 N force

so resultant force is of 4.8 N acting towards 12° north of east

alternatively, one can find components of the resultant as the sum of individual components

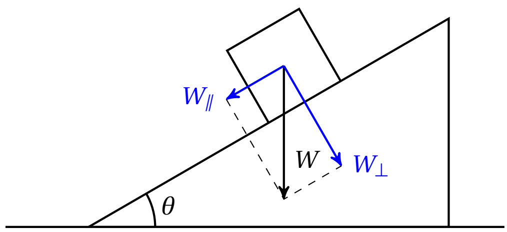
$$F_x = 3.0 + 2.0\cos 30^\circ \approx 4.73\ \mathrm{N}, \quad F_y = 2.0\sin 30^\circ = 1.0\ \mathrm{N}$$

magnitude and direction of the resultant can then be found from its components

$$F = \sqrt{F_x^2 + F_y^2} = \sqrt{4.73^2 + 1.0^2} \approx 4.84\ \mathrm{N}, \theta = \tan^{-1}\frac{F_y}{F_x} = \tan^{-1}\frac{1.0}{4.73} \approx 11.9^\circ$$

this of course agrees with scale diagram method, but resolving gives more precise results □

Example 1.7 A box of weight $W = 20.0$ N is resting on an inclined slope at 30° to the horizontal. Find the components of weight parallel to the slope and normal to the slope.



✍ the vector diagram is shown

component of weight parallel to slope: $W_{/\!/} = W\sin\theta = 20.0 \times \sin 30^\circ = 10.0$ N

component of weight normal to slope: $W_\perp = W\cos\theta = 20.0 \times \cos 30^\circ \approx 17.3$ N □

CHAPTER 2
Measurements

2.1 uncertainties

physics is a practical science, any law of physics must be evidenced by experimental facts

any meaningful physical quantity is measured either directly or indirectly

but repeated readings may not give a consistent value, instead they show a *spread* of data

uncertainty gives the *range* of values in which *true value* of the measurement is asserted to lie

measurement of a particular quantity is usually reported as $x \pm \Delta x$, where reported value x is the average of repeated readings, and Δx is its uncertainty

2.1.1 absolute uncertainty

Δx measures the size of the range of values where true value probably lies

therefore Δx is called the **absolute uncertainty**

- absolute uncertainty Δx carries same unit as quantity x
- absolute uncertainty can be worked out from *range* of readings

  range of a data set $x_1, x_2, x_3, \cdots$ is the difference between greatest and smallest value

  absolute uncertainty is given by: $\boxed{\Delta x = \frac{1}{2}(x_{\max} - x_{\min})}$
- absolute uncertainty is usually kept to one significant figure only [5]

[5] In some cases where the uncertainty of a quantity is not stated explicitly, the uncertainty is indicated by the number of significant figures in the stated value. If the height of a person is measured to be 1.75 m, this means the first two digits (1 and 7) are certain, while the last digit (5) is uncertain.

When you add or subtract numbers, the number of significant figures is determined by the location of the decimal place. For example, $1.11 + \underline{4.2} + 0.563 = 5.873$, the result should be written as 5.9. When you multiply or divided numbers, the result can have no more significant figures than the term with the fewest significant figures. For example, $1.35 \times 462 \times \underline{0.27} = 168.399$, the result should be written as 170.

However, in AS & A-Level physics, apart from the problems regarding uncertainties, it is allowed to give since Δx indicates where the readings start to get problematic

measured quantity x is kept to the same decimal place as Δx

for example, if value for the speed of an athlete is found to be $v = (8.16 \pm 0.27)$ m s^{-1}, the result, to an appropriate number of significant figures, should be kept as: $v = (8.2 \pm 0.3)$ m s^{-1}.

2.1.2 fractional & percentage uncertainty

ratio of absolute uncertainty to reported value, i.e., $\dfrac{\Delta x}{x}$, gives the **fractional uncertainty**

recording this ratio as a percentage number, this gives the **percentage uncertainty**

- fractional and percentage uncertainty have no unit
- $\dfrac{\Delta x}{x}$ gives relative measure of spread of data, so it is also called the relative uncertainty

Example 2.1 A students measures the diameter of a cylindrical bottle with a vernier calliper. The measurements are taken from several different positions and along different directions. The readings she obtained are: 4.351 cm, 4.387 cm, 4.382 cm, 4.372 cm, 4.363 cm. What is the percentage uncertainty of her measurements?

✍ average value: $d = \frac{1}{5}(4.351 + 4.387 + 4.382 + 4.372 + 4.363) = 4.371$ cm

absolute uncertainty: $\Delta d = \frac{1}{2}(d_{\text{max}} - d_{\text{min}}) = \frac{1}{2}(4.387 - 4.351) = 0.036$ cm

result of measurement should be recorded as: $d = 4.37 \pm 0.04$ cm

percentage uncertainty: $\dfrac{\Delta d}{d} = \dfrac{0.036}{4.371} \approx 0.82\%$ □

2.1.3 propagation of uncertainties

in many situations, the quantity that we want to find cannot be measured directly

the quantity of interest has to be computed from other quantities

uncertainty of this calculated quantity would depend on two things:

- uncertainties of the raw data from which it is calculated
- how the calculated quantity is related to those original quantities

suppose quantities A and B are two measurables with uncertainty ΔA and ΔB

X is a quantity to be computed by taking their sum, difference, product or quotient

---

one more significant figure that what is required. So in other sections of my notes where we do not keep track of the uncertainties, I could be a bit sloppy with the issue of significant figures when numerical values are worked out.

to evaluate uncertainty in X, we estimate the worst scenario, i.e., the greatest deviation from its reported value

addition: $S = A + B$

$$S_{\max} = A_{\max} + B_{\max} = (A+\Delta A)+(B+\Delta B) = (A+B)+(\Delta A+\Delta B) = S+(\Delta A+\Delta B) \quad \Rightarrow \quad \boxed{\Delta S = \Delta A + \Delta B}$$

subtraction: $D = A - B$

$$D_{\max} = A_{\max} - B_{\min} = (A+\Delta A)-(B-\Delta B) = (A-B)+(\Delta A+\Delta B) = D+(\Delta A+\Delta B) \quad \Rightarrow \quad \boxed{\Delta D = \Delta A + \Delta B}$$

multiplication: $P = AB$

$$P_{\max} = A_{\max} B_{\max} = (A+\Delta A)(B+\Delta B) = AB + B\Delta A + A\Delta B + \Delta A\Delta B \quad \Rightarrow \quad \Delta P = B\Delta A + A\Delta B + \Delta A\Delta B$$

divide both sides by $P = AB$, we get

$$\frac{\Delta P}{P} = \frac{\Delta A}{A} + \frac{\Delta B}{B} + \frac{\Delta A}{A} \cdot \frac{\Delta B}{B}$$

percentage uncertainty of a measurable is usually a few percent, so $\dfrac{\Delta A}{A} \cdot \dfrac{\Delta B}{B} \approx 0$

so this piece is dropped from the last expression $\quad \Rightarrow \quad \boxed{\dfrac{\Delta P}{P} = \dfrac{\Delta A}{A} + \dfrac{\Delta B}{B}}$

division: $Q = \dfrac{A}{B}$

one can show that $\boxed{\dfrac{\Delta Q}{Q} = \dfrac{\Delta A}{A} + \dfrac{\Delta B}{B}}$

the derivation is left as an exercise for the reader

power & roots: $Q = A^l B^m C^n \cdots$

percentage uncertainty in Q is: $\dfrac{\Delta Q}{Q} = l\dfrac{\Delta A}{A} + m\dfrac{\Delta B}{B} + n\dfrac{\Delta C}{C} + \cdots$

this can be thought of as a generalization for multiplication and division operations

brief summary

- for addition and subtraction, *absolute uncertainties* add up
- for multiplication, division and powers, *percentage uncertainties* add up

➢ notice that uncertainties always add

Example 2.2 The resistance of a resistor is measured. The current through the resistor is 1.8 ± 0.1 A and the potential difference across is 7.5 ± 0.2 V. What is the resistance and its uncertainty?

value of resistance: $R = \dfrac{V}{I} = \dfrac{7.5}{1.8} \approx 4.17\ \Omega$

percentage uncertainty in resistance: $\dfrac{\Delta R}{R} = \dfrac{\Delta V}{V} + \dfrac{\Delta I}{I} = \dfrac{0.2}{7.5} + \dfrac{0.1}{1.8} \approx 8.2\%$

absolute uncertainty in resistance: $\Delta R = 8.2\% \times 4.17 \approx 0.34\ \Omega$

so we find resistance of the resistor: $R = 4.2 \pm 0.3\ \Omega$ □

Example 2.3 The density of a liquid is found by measuring its mass and its volume. The following is a summary of the measurements: mass of empty beaker = (20 ± 1)g, mass of beaker and liquid = (100 ± 1)g, and volume of liquid = $(10.0 \pm 0.5)\text{cm}^3$. What is the density of this liquid and the uncertainty in this value?

✍ mass of liquid: $m = m_2 - m_1 = 100 - 20 = 80$ g

uncertainty in mass: $\Delta m = \Delta m_2 + \Delta m_1 = 1 + 1 = 2$ g

density of liquid: $\rho = \dfrac{m}{V} = \dfrac{80}{10.0} = 8.00\ \text{g cm}^{-3}$

percentage uncertainty in density: $\dfrac{\Delta \rho}{\rho} = \dfrac{\Delta m}{m} + \dfrac{\Delta V}{V} = \dfrac{2}{80} + \dfrac{0.5}{10.0} = 7.5\%$

absolute uncertainty in density: $\Delta \rho = 7.5\% \times 8.00 = 0.60\ \text{g cm}^{-3}$

so density of this liquid is recorded as: $\rho = 8.0 \pm 0.6\ \text{g cm}^{-3}$ □

Example 2.4 The period of simple pendulum is given by $T = 2\pi\sqrt{\dfrac{L}{g}}$. In an experiment, the length of string is measured to be $L = 100.0 \pm 0.5$ cm, and the time taken for 10 full oscillations is $t = 20.0 \pm 0.2$ s. What is the value for acceleration of free fall g and its uncertainty?

✍ period of one oscillation: $T = \dfrac{1}{10} t = 2.00 \pm 0.02$ s

let's rearrange $T = 2\pi\sqrt{\dfrac{L}{g}}$ into $g = \dfrac{4\pi^2 L}{T^2}$

acceleration of free fall: $g = \dfrac{4\pi^2 \times 100.0}{2.00^2} \approx 987\ \text{cm s}^{-2}$

percentage uncertainty: $\dfrac{\Delta g}{g} = \dfrac{\Delta L}{L} + 2\dfrac{\Delta T}{T} = \dfrac{0.5}{100.0} + 2 \times \dfrac{0.02}{2.00} = 2.5\%$ [6]

absolute uncertainty: $\Delta g = 2.5\% \times 987 \approx 24.7\ \text{cm s}^{-2}$

so result of this measurement is: $g = 990 \pm 20\ \text{cm s}^{-2}$ □

[6] Starting from the formula $T = 2\pi\sqrt{\dfrac{L}{g}}$, it is attempting to write $\dfrac{\Delta T}{T} = \dfrac{1}{2}\dfrac{\Delta L}{L} + \dfrac{1}{2}\dfrac{\Delta g}{g}$. But this would mean that T is a calculated quantity whose uncertainty is determined by the uncertainty in L and the uncertainty in g, which is incorrect. The right way to do it is to rearrange the formula so that calculated quantity of interest is the subject of the working equation, the propagation of uncertainties then becomes explicit.

2.2 errors of measurement

difference between the measured value and the true value is called **error**

total error is usually a combination of two components: systematic error and random error

2.2.1 systematic & random errors

> **systematic errors** cause all readings to be greater or smaller than the true value by the same amount

- faulty equipments, biased observers, calibration errors could produce systematic errors

  examples of systematic errors include:
  - a vernier calliper does not read zero when fully closed, this introduces **zero error**
  - one always reads a measuring cylinder from a higher angle, this introduces **parallax error**
  - spring of force meter becomes weaker over time, so force meter always gives overestimates
- systematic errors can be reduced by using better equipments or methods
  - one can check for zero error before taking readings with a micrometer
  - one can *calibrate* a balance with a known mass before using it to measure mass of an object

> **random errors** cause repeated readings to fluctuate above or below the actual value

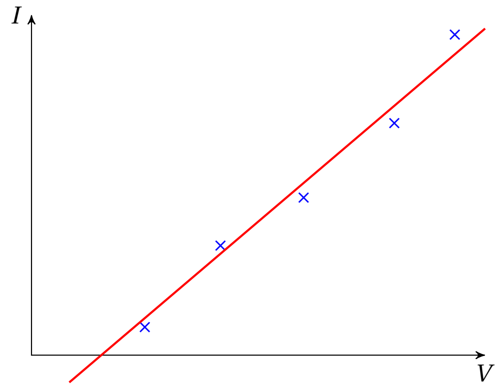
- deviations caused by random error are unpredictable
- insensitive equipments, lack of observer precision, changes in environment, imprecise definitions could produce random errors

  examples of causes of random errors include:
  - **human reaction errors** when measuring a time quantity on a stop-watch
  - electronic noise due to thermal vibrations of atoms when measuring an electric current
  - when measuring length of a crack, different people could pick different end points
- random errors can be reduced by averaging the results from repeated measurements

  for example, diameter of a sphere can be measured along different directions and averaged
- random errors can also be reduced by using better equipment and better technique

  for example, time for an object to fall can be measured with a light gate, instead of a stopwatch

Example 2.5 An experiment is carried out to measure the resistance of a metallic resistor, which is known to be constant throughout the experiment. A set of readings for voltage V across the resistor and the corresponding current I are obtained. A graph of I against V is plotted as shown. What can you say about the errors of the experiment?



✍ one can first draw a *best fit line* to see the distribution of data points

constant resistance means I should be directly proportional to V

so the best fit should be a straight line through the origin

but the best fit does not pass through origin, so there exists systematic error

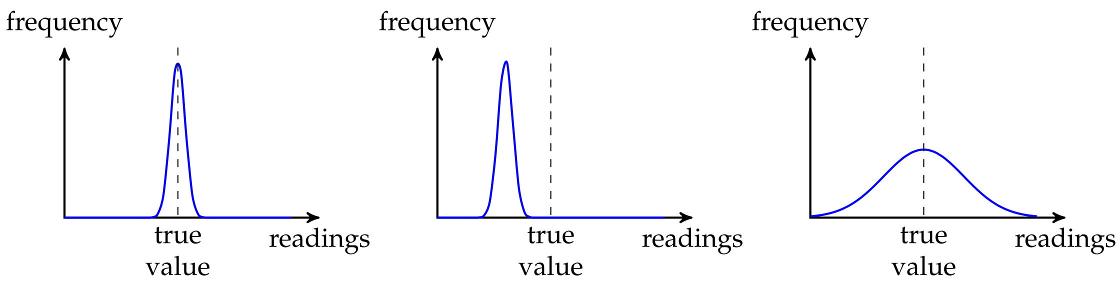
also data points scatter above and below the best fit, so random errors are present □

2.2.2 accuracy & precision

to analyse the result of an experiment, two important aspects are accuracy and precision

measurement is said to be **accurate** if the result is close to the true value

measurement is said to be **precise** if repeated readings are close to each other



(a) precise and accurate (b) precise but not accurate (c) accurate but not precise

distribution of readings with different precision and accuracy

➢ accuracy of a measurement is closely related to systematic errors

large systematic errors mean the results must be inaccurate

- precision of a measurement is closely related to random errors

  large random errors cause repeated readings to spread, so the result must be imprecise
- precision is usually indicated by the percentage uncertainty of the measurement

  similarly, precision is also indicated by the number of significant figures in a measurement

  for example, metre rule has a precision of 0.1 cm, while micrometer has a precision of 0.001 cm

Example 2.6 The value for the acceleration of free fall is determined in an experiment. The result is reported to be $g = 14 \pm 5$ m s^{-2}. Is this result accurate? Is it precise?

✍ true value for g is around 9.8 m s^{-2}

stated value is not close to the true value, so the result is not accurate

percentage uncertainty in this result is $\frac{5}{14} \approx 36\%$, which is quite large

so the result is not precise either □

CHAPTER 3

Kinematics

Kinematics is the study of motion. We will introduce three kinematic quantities, displacement, velocity and acceleration, and see how these terms are used to describe an object's motion.

3.1 kinematic quantities

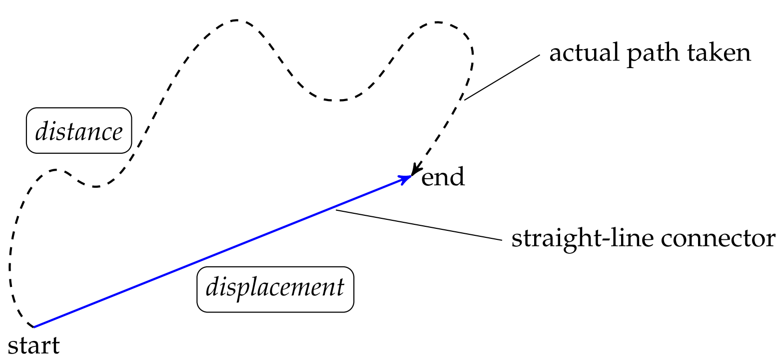
3.1.1 displacement & distance

in everyday language, we talk about the **distance** travelled by an object, which usually refers to the length travelled by an object without considering in which direction it moves

to fully describe position of an object, we also need specify where it is heading

> **displacement** is defined as the distance moved by an object in a specific direction

- displacement and distance are measured in metres, or any reasonable length units
- displacement is a vector quantity, while distance is a scalar
- displacement is the *straight-line* distance pointing from starting point towards end point
  even if actual path taken is curved, displacement is always the straight-line distance



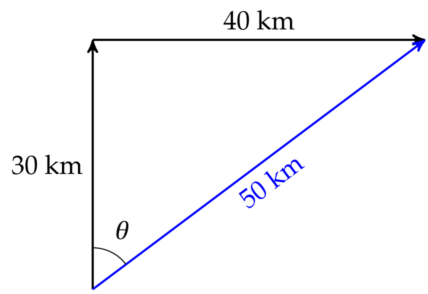
difference between displacement and distance

Example 3.1 An athlete is running around a circular track of radius 60 m. When he completes one lap, what is the distance moved out? What about his displacement?

✍ distance moved is the perimeter of the circle: $s = 2\pi r = 2\pi \times 60 \approx 380$ m

athlete returns to same starting point after one lap, so displacement is zero □

Example 3.2 A ship travels 30 km north, takes a right, and then travels 40 km east to reach its destination. Compare the distance and the displacement travelled.



✍ sum of all lengths gives distance: $30 + 40 = 70$ km

displacement vector is shown on the graph

magnitude of displacement = $\sqrt{30^2 + 40^2} = 50$ km

it is at an angle of $\theta = \tan^{-1}\dfrac{40}{30} \approx 53^\circ$ east of north □

3.1.2 velocity & speed

displacement of a moving object may change with respect to time

an object is moving fast if it has a large change in displacement during a given time interval

> change in displacement per unit time is called the **velocity** of the object: $\boxed{v = \dfrac{\Delta s}{\Delta t}}$

- SI unit of measurement for velocity: $[v] =$ m s^{-1}
- velocity is a vector

  this follows from the fact that displacement is a vector quantity
- for *linear* motion, one shall pick a specific direction as the positive direction

  then a negative velocity implies motion in the opposite direction
- defining equation $v = \dfrac{\Delta s}{\Delta t}$ gives the *average* value for velocity or speed over an interval Δt

  more precisely: $\boxed{\text{average velocity} = \dfrac{\text{total displacement}}{\text{time taken}}}$, and $\boxed{\text{average speed} = \dfrac{\text{total distance}}{\text{time taken}}}$

  this should be distinguished from the notion of *instantaneous velocity*

  instantaneous velocity is defined as the rate of change in displacement at a particular instant

  if we take a very short interval Δt, as $\Delta t \to 0$, average velocity tends to instantaneous velocity

  this is expressed in a compact differential form: $v = \lim\limits_{\Delta t \to 0} \dfrac{\Delta s}{\Delta t} \Rightarrow \boxed{v = \dfrac{\mathrm{d}s}{\mathrm{d}t}}$

➢ it is also common to use **speed** to describe how fast an object moves

speed is defined as the change of the distance travelled per unit time

velocity can be thought as speed in a particular direction

Example 3.3 A cyclist travels a distance of 3.0 km in 20 minutes. She rests for 15 minutes. She then covers a further distance of 5.1 km in a time of 40 minutes. Calculate the average speed of the cyclist in m s^{-1}: (a) during the first 20 minutes of the journey, (b) for the whole journey.

✍ for the first 20 minutes: $v = \dfrac{3.0 \times 10^3}{20 \times 60} = 2.5\ \mathrm{m\ s^{-1}}$

for whole journey: $v = \dfrac{(3.0 + 0 + 5.1) \times 10^3}{(20 + 15 + 40) \times 60} = 1.8\ \mathrm{m\ s^{-1}}$ □

Example 3.4 A man walks along a straight road for a distance of 800 m in 5.0 minutes. He then turns around, and walks along the same road for a distance of 280 m in 3.0 minutes. What is the average speed and the average velocity of this man during the 8.0 minutes?

✍ total distance travelled = 800 + 280 = 1080m, so average speed: $v = \dfrac{1080}{8.0 \times 60} = 2.25\ \mathrm{m\ s^{-1}}$

change of displacement = 800 + (−280) = 520m, so average velocity: $v = \dfrac{520}{8.0 \times 60} \approx 1.08\ \mathrm{m\ s^{-1}}$ □

Example 3.5 A maglev train travels at an average speed of 60 m s^{-1} from the city centre to the airport, and at 40 m s^{-1} on its return journey over the same distance. What is the average speed of the round-trip? What about the average velocity?

✍ suppose the distance between airport and city centre is S

average speed: $v = \dfrac{2S}{t_1 + t_2} = \dfrac{2S}{\frac{S}{60} + \frac{S}{40}} = 48\ \mathrm{m\ s^{-1}}$

for a round-trip, train returns to same staring position

change in displacement is zero, so average velocity is zero □

3.1.3 acceleration

velocity of a moving object may change, i.e., objects can speed up, slow down or make turns

> change in velocity per unit time is defined as the **acceleration**: $\boxed{a = \dfrac{\Delta v}{\Delta t}}$

➢ unit of measurement for acceleration: $[a] = \mathrm{m\ s^{-2}}$

➢ acceleration is a vector quantity, it has both magnitude and direction

this is because of vector nature of velocity, change in velocity must also have direction

➢ for *linear* motion, one usually pick direction of initial velocity as positive direction

$a > 0$ would imply acceleration in the normal sense, i.e., motion with an increasing speed

$a < 0$ would imply deceleration, i.e., motion with a decreasing speed

➢ when velocity changes, it could be change in magnitude or/and change in direction [7]

for example, for an object moving along a curved path, its velocity is constantly changing direction, so it must have a non-zero acceleration

no acceleration would imply no change in speed and no change in direction of motion

➢ defining equation $a = \dfrac{\Delta v}{\Delta t}$ gives average acceleration over time interval Δt

we can likewise introduce *instantaneous acceleration* as the rate of change in velocity

taking the limit where the time interval $\Delta t \to 0$. we have: $a = \lim\limits_{\Delta t \to 0} \dfrac{\Delta v}{\Delta t} \Rightarrow \boxed{a = \dfrac{\mathrm{d}v}{\mathrm{d}t}}$

Example 3.6 A ball hits a barrier at right angles with a speed of 15 m s^{-1}. It makes contact with the barrier for 30 ms and then rebounds with a speed of 12 m s^{-1}. What is the average acceleration during the time of contact?

✍ note that direction of velocity changed during rebound, so $\Delta v = 15 - (-12) = 27$ m s^{-1}

average acceleration: $a = \dfrac{\Delta v}{\Delta t} = \dfrac{27}{30 \times 10^{-3}} = 900$ m s^{-2} □

3.2 motion graphs

how one quantity changes with another quantity can be visually shown on a *graph*

changes in displacement, velocity or acceleration over time can be shown on *motion graphs*

as we will see, s-t graph, v-t graph and a-t graphs are closely interrelated to one another

3.2.1 displacement-time graphs

a displacement-time graph shows an object's position at any given time

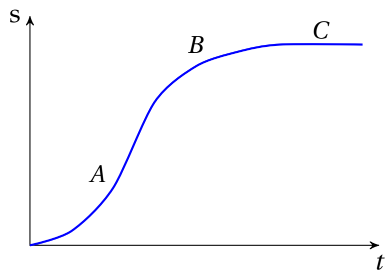
➢ gradient of tangent gives rate of change in displacement

but this is instantaneous velocity, which can be given by $v = \dfrac{\mathrm{d}s}{\mathrm{d}t}$, so we have:

$$\boxed{\text{velocity} = \text{gradient of } s\text{-}t \text{ graph}}$$

[7] Acceleration of an object can be considered as the combination of two components. One component is known as the *normal* acceleration or the *centripetal* acceleration, which is at right angle to the velocity and is responsible for the change in direction of motion. The other component is called the *tangential* acceleration, which is parallel to the direction of motion and causes change in magnitude of object's velocity. You will learn more about these in further mechanics.

Example 3.7 Describing the motion from the displacement-time graph shown.

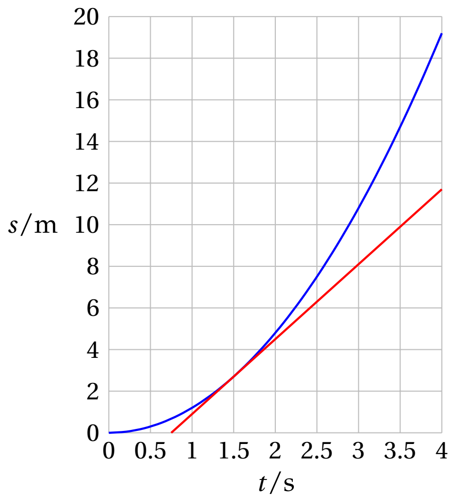


stage A: gradient of the graph is increasing, showing that the object is speeding up

stage B: gradient starts to decrease, so the object gradually slows down

stage C: curve becomes horizontal, gradient becomes zero, means that the object eventually comes to a stop □

Example 3.8 The diagram shows the displacement-time graph for a vehicle travelling along a straight road. Use the graph to find, (a) the average velocity during the first 4.0 s of the motion, (b) the velocity of the vehicle *at* time $t = 1.5$ s.



✍ during first 4.0 s, average velocity is

$$v = \frac{\Delta s}{\Delta t} = \frac{19.2}{4.0} \approx 4.8 \text{ m s}^{-1}$$

to find velocity at $t = 1.5$ s, a tangent is drawn

gradient of tangent gives instantaneous velocity:

$$v = \frac{11.6 - 0}{4.0 - 0.75} \approx 3.6 \text{ m s}^{-1}$$ □

3.2.2 velocity-time graphs

a velocity-time graph shows the velocity of a moving object at any instant

➢ since the rate of change of velocity gives the acceleration, so

$$\boxed{\text{acceleration} = \text{gradient of } v\text{-}t \text{ graph}}$$

➢ v-t graph also gives information about the change in displacement, that is

$$\boxed{\text{change in displacement} = \text{area under } v\text{-}t \text{ graph}}$$

reasoning: in very short interval Δt_i, change in velocity is small so $v(t_i) \approx$ constant during Δt_i

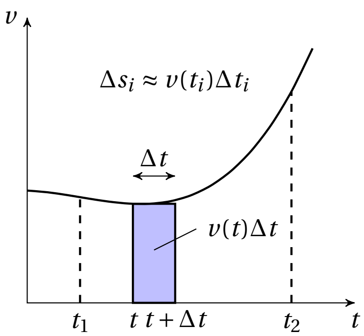
displacement moved out $\Delta s_i \approx v(t_i)\Delta t_i$, which corresponds to area of a thin rectangle

sum of all these small Δs_i's gives total change in displacement over a period of time

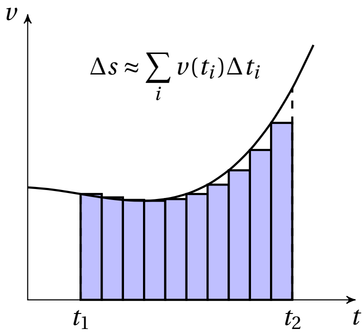
now consider the limit where each of the time interval $\Delta t_i \to 0$

total area of these rectangles approximates area bounded by the v-t curve and time axis [8]

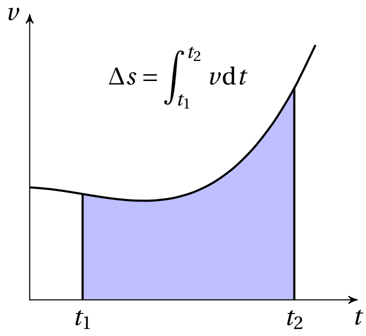
[8]Mathematically, integration is the inverse operation of taking derivatives. By definition $v = \frac{\mathrm{d}s}{\mathrm{d}t}$, then it follows naturally that $\Delta s = \int v\,\mathrm{d}t$. While the derivative of a given function gives the gradient of tangent at



(a) displacement Δs in short interval Δt



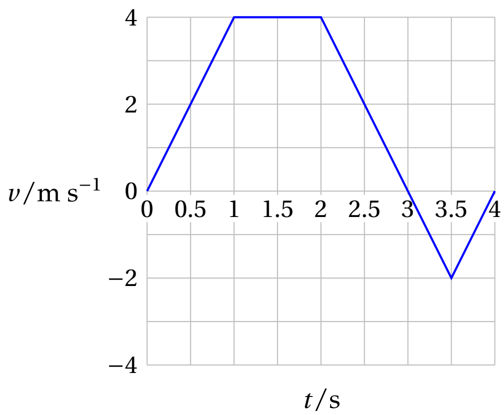
(b) total displacement estimated by summing the many Δs_i's



(c) total displacement as area under v-t graph

calculating change in displacement by finding the area under velocity-time graph

Example 3.9 The velocity of a toy car is shown. For the journey shown on the graph, use the graph to find (a) the total distance travelled, and (b) the total displacement travelled.



✍ distance is estimated using area under v-t graph

0~3.0 s: $s_1 = \frac{1}{2} \times 1.0 + 3.0 \times 4.0 = 8.0$ m

3.0~4.0 s: $s_2 = \frac{1}{2} \times 1.0 \times 2.0 = 1.0$ m

total distance $= 8.0 + 1.0 = 9.0$ m

total displacement $= (+8.0) + (-1.0) = 7.0$ m □

3.2.3 acceleration-time graphs

one can similarly plot an acceleration-time graph to give the changes in acceleration

➢ a-t graphs can give information about changes in velocity

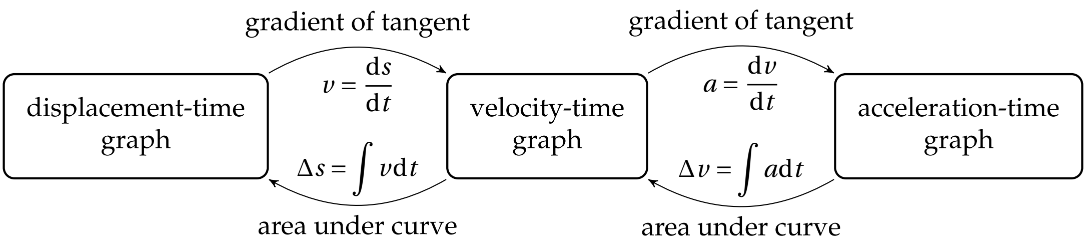
similar discussions lead to the following conclusion:[9]

$$\boxed{\text{change in velocity} = \text{area under } a\text{-}t \text{ graph}}$$

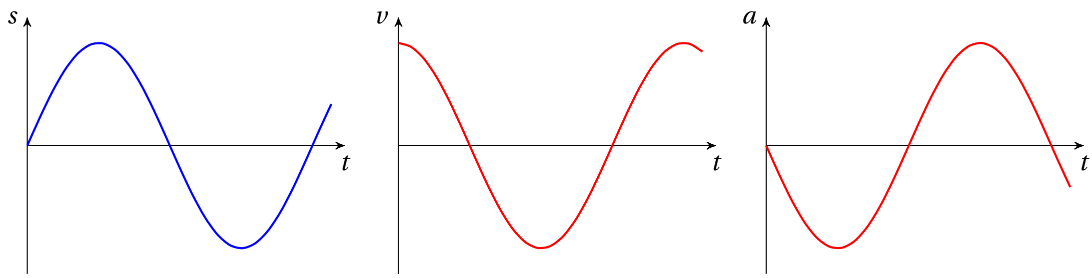
each point on its graph, integrating a function gives the signed area bounded by the graph. The reader may find the formal treatment of this relationship in any calculus textbook.

[9]Using area under a-t graph to find changes in velocity is not required in the AS-Level physics syllabus. I am putting this in the notes mainly for the completeness of the discussions on motion graphs.

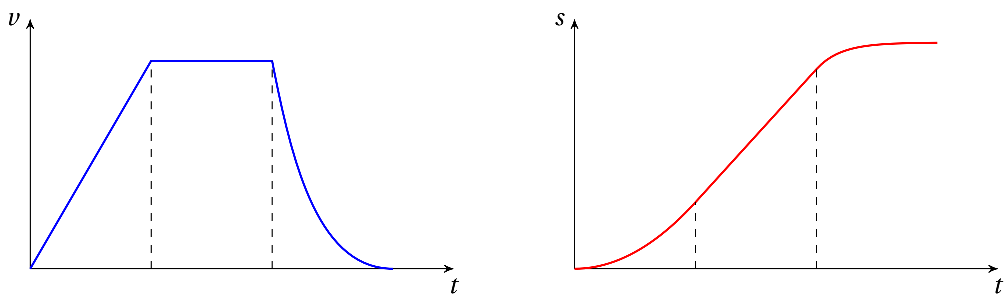
relationships between displacement, velocity and acceleration graphs are summarised below



Example 3.10 Given the displacement-time graph as shown, check yourself that this s-t graph leads to the velocity-time graph and the acceleration-time graph shown.



Example 3.11 Given the velocity-time graph as shown, check yourself that this v-t graph leads to the displacement-time graph as shown.



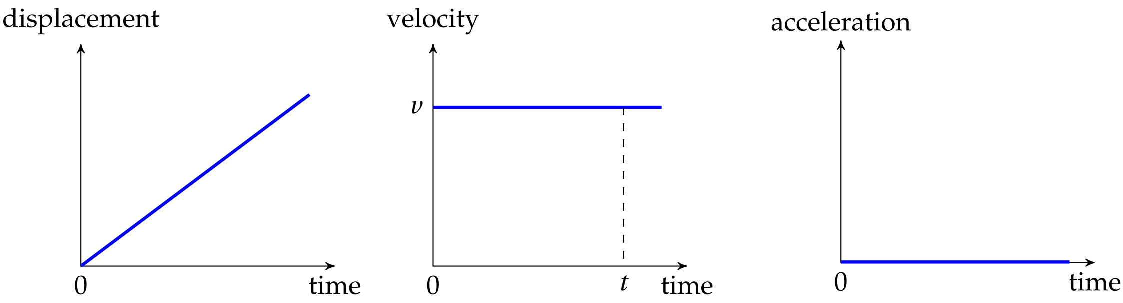
3.3 linear motion with constant velocity

let's look at the simplest kind of motion

that is, an object moving at constant speed in a straight line: $v = \text{constant}$

the equations of motion are straightforward: $\boxed{a = 0}$ $\boxed{s = vt}$ [10]

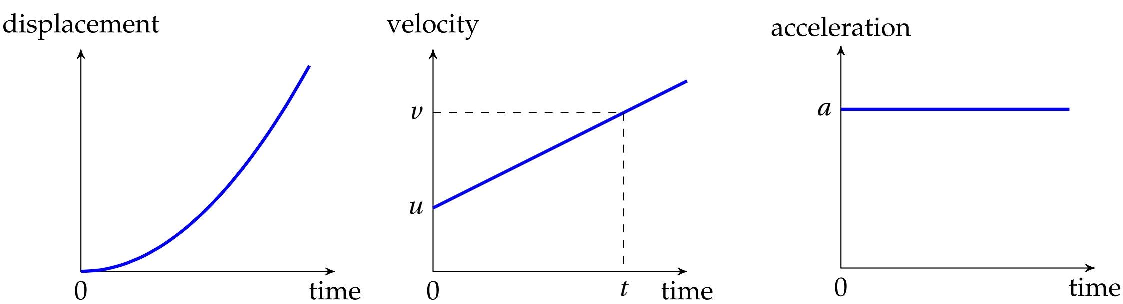
[10] It is implicitly assumed that the motion starts from the origin with respect to which displacement is defined. More generically, we should write: $s = s_0 + vt$, where s_0 is the initial displacement.



motion graphs for linear motion at constant velocity

3.4 linear motion with constant acceleration

the second simplest type of motion is a linear motion with acceleration $a = \text{constant}$



motion graphs for linear motion at uniform acceleration

3.4.1 equations of motion

during a time interval t, suppose velocity changes from initial value u to final value v

from the defining equation of acceleration $a = \frac{\Delta v}{\Delta t} = \frac{v-u}{t}$, we get

$$\boxed{v = u + at}\text{[11]} \tag{3.1}$$

to find total displacement travelled, we compute the area under the v-t graph

$$\boxed{s = \frac{1}{2}(u+v)t} \tag{3.2}$$

for which we can interpret $\bar{v} = \frac{1}{2}(u+v)$ as the average velocity during that time

[11]Proof with calculus: $\mathrm{d}v = a\mathrm{d}t \Rightarrow \Delta v = \int_u^v \mathrm{d}v = \int_0^t a\mathrm{d}t \Rightarrow v - u = at \Rightarrow v = u + at$

plug (3.1) into (3.2), we find an expression for the displacement travelled in terms of u and a:

$$\boxed{s = ut + \frac{1}{2}at^2} \tag{3.3}$$ [12][13]

this shows the displacement s is a quadratic function in time t

this is consistent with the parabolic shape of the s-t graph shown

we can also eliminate the time variable t to derive one last equation

from (3.1) we have $t = \dfrac{v-u}{a}$, substitute this into (3.2), we find

$$s = \frac{1}{2}(u+v) \times \frac{v-u}{a} \quad \Rightarrow \quad \boxed{2as = v^2 - u^2} \tag{3.4}$$

Example 3.12 A car starts from rest and accelerates uniformly at 5.0 m s^{-2} for 6.0 s. (a) How fast is the car travelling at $t = 8.0$ s? (b) What is the distance travelled by the car in this time?

✍
$$v = u + at \quad \Rightarrow \quad v = 0 + 5.0 \times 6.0 = 30 \text{ m s}^{-1}$$
$$s = ut + \frac{1}{2}at^2 \quad \Rightarrow \quad s = 0 + \frac{1}{2} \times 5.0 \times 6.0^2 = 90 \text{ m}$$ □

Example 3.13 A car is travelling at 30 m s^{-1}. A hazard appears in front of the car, and the driver takes immediate action to stop the car. When brakes are applied, deceleration of the car is 5.0 m s^{-2}. What is the braking distance?

✍
$$2as = v^2 - u^2 \quad \Rightarrow \quad s = \frac{v^2 - u^2}{2a} = \frac{0^2 - 30^2}{2 \times (-5.0)} = 90 \text{ m}$$ □

Example 3.14 At the instant the traffic light turns green, a motorcycle waiting at the stop line starts with a constant acceleration of 2.0 m s^{-2}. At the same instant, a truck at a constant speed of 16 m s^{-1} overtakes and passes the motorcycle. How far beyond the stop line will the motorcycle overtake the truck?

✍ suppose overtake occurs at time t after motorcycle starts to accelerate

distance travelled by motorcycle: $s_m = u_m t + \frac{1}{2}at^2 \quad \Rightarrow \quad s_m = 0 + \frac{1}{2} \times 2.0 \times t^2$

distance travelled by truck: $s_t = v_t t \quad \Rightarrow \quad s_t = 16t$

overtake when $s_m = s_t \quad \Rightarrow \quad \frac{1}{2} \times 2.0 \times t^2 = 16t \quad \Rightarrow \quad t = 16$ s

substitute t into either s_m or s_t, one finds distance travelled: $s = 256$ m □

[12] Proof with calculus: $\mathrm{d}s = v\mathrm{d}t \Rightarrow \Delta s = \int_0^s \mathrm{d}s = \int_0^t v\mathrm{d}t \Rightarrow s = \int_0^t (u + at)\mathrm{d}t = \left(ut + \frac{1}{2}at^2\right)\Big|_0^t \Rightarrow s = ut + \frac{1}{2}at^2$

[13] Equation (3.3) assumes a zero initial displacement at $t = 0$. If there is a non-zero initial displacement, one should write $s = s_0 + ut + \frac{1}{2}at^2$. Similar discussion applies to equation (3.2).

3.4.2 free fall

a typical example of uniformly accelerated motion is the free fall

everything has the tendency to fall towards ground due to earth's gravity

in this section, we assume that effects of air resistance are negligible

acceleration of free fall is then constant $a = g$

value of g does not depend on mass of the falling object[14]

➢ near surface of earth, value of acceleration of free fall: $g \approx 9.81 \text{ m s}^{-2}$

value of g could be different on a different planet

➢ for a freely-falling object released from rest, its velocity increases with time as

$$v = u + at = 0 + gt \quad \Rightarrow \quad v = gt$$

the distance it has fallen from the point of release is

$$s = ut + \frac{1}{2}at^2 = 0 \cdot t + \frac{1}{2}gt^2 \quad \Rightarrow \quad s = \frac{1}{2}gt^2$$

Example 3.15 An object is released from rest from a height of $h = 24$ m and falls freely under gravity. Air resistance is negligible. (a) How long does it take to hit the ground? (b) What is its speed when hitting the ground?

✍

$$h = \frac{1}{2}gt^2 \quad \Rightarrow \quad t = \sqrt{\frac{2h}{g}} = \sqrt{\frac{2 \times 24}{9.81}} \approx 2.21 \text{ s}$$

$$v = gt = 9.81 \times 2.21 \quad \Rightarrow \quad v \approx 21.7 \text{ m s}^{-1}$$

to find final velocity, one can also use $v^2 = u^2 + 2as$:

$$v^2 = \cancelto{0}{u^2} + 2gh = 2 \times 9.81 \times 24 \quad \Rightarrow \quad v \approx 21.7 \text{ m s}^{-1}$$ □

Example 3.16 A photograph is taken for a small particle falling from rest. The photograph is taken at 0.400 s after the object is released. Since the particle is still moving when the photograph is being taken, the image is blurred. It is found the blurred part corresponds to a length of 20.8 cm moved out by the particle, what is time of exposure for the photograph?

✍ from $t = 0$ to right before photo is taken:

$$s_1 = \frac{1}{2}gt_1^2 = \frac{1}{2} \times 9.81 \times 0.400^2 \approx 0.785 \text{ m}$$

from $t = 0$ to right after photo has been taken:

$$s_2 = s_1 + \Delta s = \frac{1}{2}gt_2^2 \quad \Rightarrow \quad t_2 = \sqrt{\frac{2(s_1 + \Delta s)}{g}} = \sqrt{\frac{2 \times (0.785 + 0.208)}{9.81}} \approx 0.450 \text{ s}$$

time of exposure: $\Delta t = t_2 - t_1 = 0.450 - 0.400 \approx 0.050$ s □

[14]The reason for this constant acceleration of free fall will be elaborated in §4.3.1.

3.4.3 upward projection

like a freely-falling object, an object tossed upwards experiences the same constant downward acceleration $a = g \approx 9.81$ m s^{-2} as long as resistive forces can be ignored

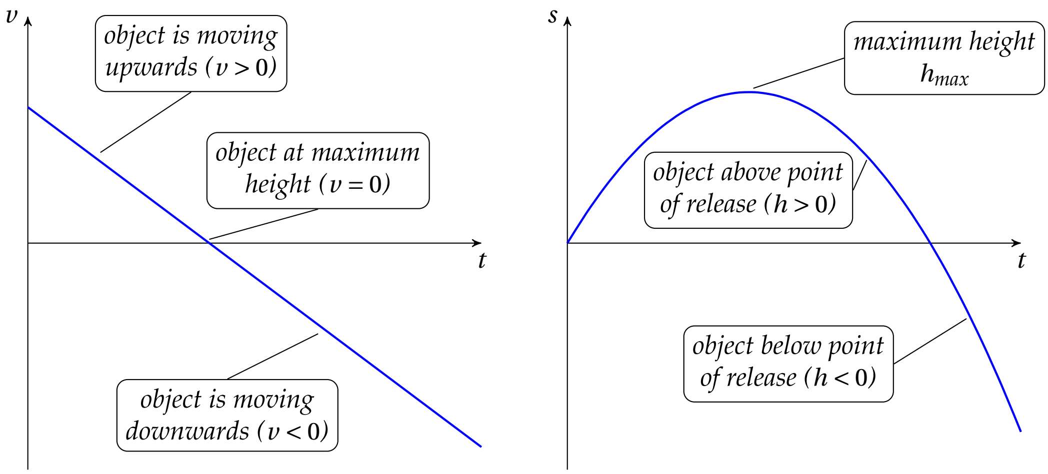
note that initial velocity u is upwards, but acceleration a is downwards so we will have different signs for u and a in the equations

conventionally, positive direction is taken as same direction as initial velocity

in our case, positive direction is upwards, the acceleration is then negative: $a = -g$

so the velocity-time relation and displacement-time relation are

$$v = u - gt \qquad s = ut - \frac{1}{2}gt^2$$



v-t graph and s-t graph for upward projectile motion

- sign of v now tells direction of motion

  $v > 0$ means object is moving upwards, $v < 0$ means it has reversed direction and starts falling

  in particular, object attains greatest height when $v = 0$

- sign of s gives whether object is at a higher or lower position with respect to point of release

  $s > 0$ means the object is above the position from which it is projected

  $s < 0$ means it is now below the point of release

Example 3.17 A ball is projected vertically upwards at 12 m s^{-1}. Air resistance is negligible. (a) Find the time taken for the ball to reach the highest position. (b) Find the greatest height.

✍ maximum height is reached when $v = 0$, so

$$v = u - gt = 0 \quad \Rightarrow \quad t = \frac{u}{g} = \frac{12}{9.81} \approx 1.22 \text{ s}$$

$$H_{\max} = ut - \frac{1}{2}gt^2 = 12 \times 1.22 - \frac{1}{2} \times 9.81 \times 1.22^2 \approx 7.34 \text{ m}$$

it is also possible to use $v^2 - u^2 = 2as$ to find $H_{\max}$, this is:

$$0^2 - u^2 = 2(-g)H_{\max} \quad \Rightarrow \quad H_{\max} = \frac{u^2}{2g} = \frac{12^2}{2 \times 9.81} \approx 7.34 \text{ m}$$ □

Example 3.18 A stone is thrown vertically upwards with an initial velocity of 14.0 m s^{-1} from the edge of a cliff that is 35 m from the sea below. (a) Find the speed at which it hits the sea. (b) Find the time taken for the stone to hit the sea.

✍ take positive direction to point upwards, we use $v^2 - u^2 = 2as$ to find

$$v^2 = 14.0^2 + 2 \times (-9.81) \times (-35) \approx 883 \text{ m}^2 \text{ s}^{-2} \quad \Rightarrow \quad v \approx -29.7 \text{ m s}^{-1}$$[15]

to find time, we can use $v = u - gt$, hence: $t = \dfrac{v-u}{-g} = \dfrac{-29.7 - 14.0}{-9.81} \approx 4.46$ s

one can also attempt $s = ut - \frac{1}{2}gt^2$, this leads to the equation: $-35 = 14.0t - \frac{1}{2} \times 9.81t^2$

this quadratic equation in t gives two roots: $t_1 \approx 4.46$ s, and $t_2 \approx -1.60$ s

negative root should be discarded since it means stones hits the sea before it is thrown

so time taken for stone to hit the sea is $t \approx 4.46$ s □

3.5 motion in two dimensions – projectile motion

a **projectile** is an object whose motion is only affected by gravity

for projectile motion, we assume no air resistance and no other forces

gravity causes a constant acceleration of free fall that acts vertically downwards

➢ curved path of a projectile is the combination of its *horizontal* and *vertical* motion

– horizontally: no acceleration, so horizontal component of velocity v_x = constant

– vertically: constant acceleration, vertical component of velocity v_y varies over time

as a consequence, a projectile would follows a *parabolic* path as it travels[16]

let's consider a projectile launched at initial velocity u at angle θ to the horizontal

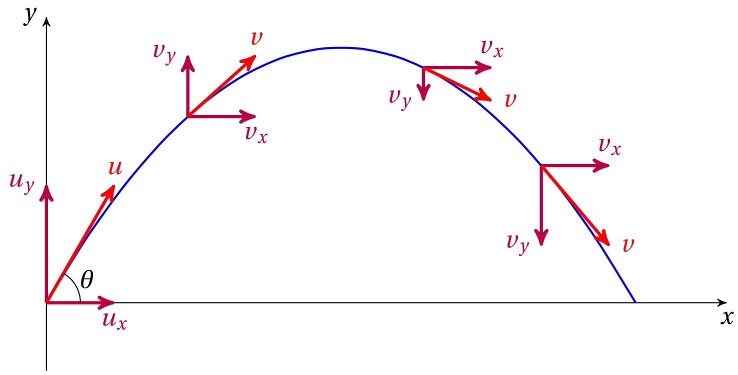
➢ horizontally, projectile maintains a constant velocity, so

$$v_x = u_x \qquad x = u_x t$$

where $u_x = u\cos\theta$ is horizontal component of initial velocity

[15]Note that we have substituted $a = -g$ since acceleration of free fall always points downwards, and $s = -35$ m since sea is below point of release. Also final velocity when hitting water is downwards, which should take a negative sign, so we discarded the positive solution for v.

[16]You may try to prove this statement with a bit algebra.



components of the velocity of a projectile at different points along its path

➢ vertically, if upward direction is taken to be positive, then acceleration $a = -g$, so
$$v_y = u_y + at = u_y - gt \qquad y = u_y t + \frac{1}{2}at^2 = u_y t - \frac{1}{2}gt^2$$
where $u_y = u\sin\theta$ is vertical component of initial velocity

➢ components can be combined to give resultant velocity or resultant displacement:
$$v = \sqrt{v_x^2 + v_y^2} \qquad s = \sqrt{x^2 + y^2}$$

maximum height reached by an projectile

when a projectile reaches the highest position, its instantaneous vertical velocity becomes zero

we can then find the time it takes to attain this maximum height.
$$v_y = u\sin\theta - gt = 0 \quad \Rightarrow \quad t = \frac{u\sin\theta}{g}$$
to find $H_{\max}$, one can use either equation (3.2) or (3.3)
$$H_{\max} = \frac{1}{2}(u_y + v_y)t = \frac{1}{2}u_y t = \frac{1}{2} \times u\sin\theta \times \frac{u\sin\theta}{g} = \frac{u^2\sin^2\theta}{2g}$$
$$H_{\max} = u_y t - \frac{1}{2}gt^2 = u\sin\theta \times \frac{u\sin\theta}{g} - \frac{g}{2} \times \left(\frac{u\sin\theta}{g}\right)^2 = \frac{u^2\sin^2\theta}{2g}$$
➢ for the same initial speed u, the greater the angle of projection, the higher the object can get

in the extremal case where $\theta = 90^\circ$, it simply becomes an upward projection motion

airborne time and horizontal range of an projectile

a ball projected from the ground will first rise in height

but it will eventually fall to the ground due to the gravitational pull after a period of time T

when it lands, its vertical displacement is zero, so
$$Y = u_y T - \frac{1}{2}gT^2 = u\sin\theta T - \frac{1}{2}gT^2 = 0 \quad \Rightarrow \quad T = \frac{2u\sin\theta}{g}$$

the horizontal range is given by

$$X = u_x T = u\cos\theta \times \frac{2u\sin\theta}{g} = \frac{2u^2\sin\theta\cos\theta}{g} \quad \Rightarrow \quad X = \frac{u^2\sin 2\theta}{g}$$

where in the last step the trigonometric identity $\sin 2\alpha = 2\sin\alpha\cos\alpha$ has been used

➢ for same initial speed u, projectile launcher at greater angle stays in air for longer time

greatest airborne time is obtained if object is projected straight up, i.e., $\theta = 90^\circ$

➢ for same initial speed, horizontal range of projectile depends on angle θ of projection

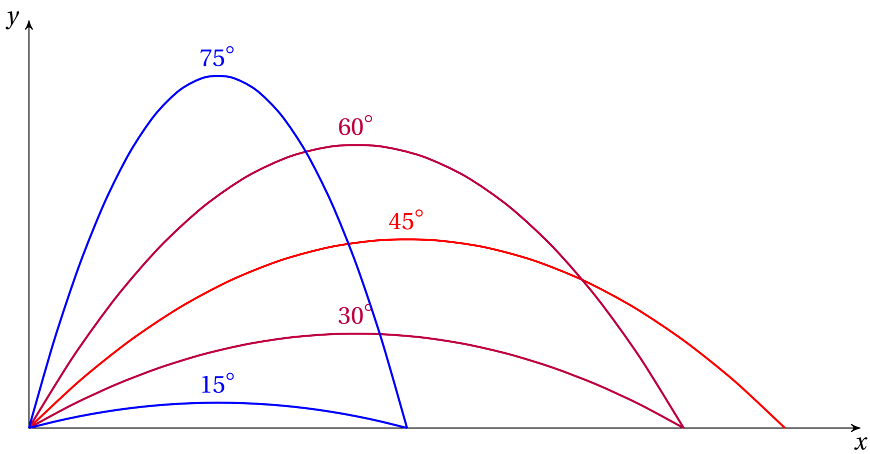
to obtain the greatest horizontal range, two things are required

- sufficiently large horizontal velocity v_x
- sufficiently long time T staying in the air

however, a larger v_x requires a smaller θ, which means a shorter airborne time T

therefore, there is a compromise between the two

optimal angle should be neither be too large nor too small, which can be shown to be 45°



trajectories of projectiles launched at the same speed but different angles

Example 3.19 A ball is thrown from a point O at 15 m s^{-1} at an angle of 40° to the horizontal. The ball reaches its highest position at point P. Ignore the effects of air resistance. (a) How long does it take to reach P? (b) What is the magnitude of the displacement OP?

✍ at highest point: $v_y = u_y - gt = 0 \quad \Rightarrow \quad t = \dfrac{u\sin\theta}{g} = \dfrac{15\times\sin 40^\circ}{9.81} \approx 0.983$ s

vertical displacement: $y = u_y t - \frac{1}{2}gt^2 = 15\sin 40^\circ \times 0.983 - \frac{1}{2}\times 9.81 \times 0.983^2 \approx 4.74$ m

horizontal displacement: $x = u_x t = 15\cos 40^\circ \times 0.983 \approx 11.3$ m

resultant displacement: $|OP| = \sqrt{x^2+y^2} = \sqrt{11.3^2 + 4.74^2} \approx 12.2$ m □

Example 3.20 A small object is horizontally projected at 7.20 ms^{-1} from a surface at a height of $h = 1.2$ m above the ground. Assume there is no air resistance. (a) What is the time taken for the object to hit the ground? (b) What is the horizontal range? (c) Find the velocity at which the object hits the ground.

✍ vertically, take downward as positive: $h = \cancelto{0}{u_y t} + \frac{1}{2} g t^2 \Rightarrow t = \sqrt{\frac{2h}{g}} = \sqrt{\frac{2 \times 1.2}{9.81}} \approx 0.495$ s

horizontal range: $x = u_x t = 7.20 \times 0.495 \approx 3.56$ m

final vertical velocity: $v_y = \cancelto{0}{u_y} + gt = 9.81 \times 0.495 \approx 4.85 \text{ m s}^{-1}$

magnitude of resultant velocity: $v = \sqrt{v_x^2 + v_y^2} = \sqrt{7.20^2 + 4.85^2} \approx 8.68 \text{ m s}^{-1}$

angle to which resultant velocity makes with horizontal: $\phi = \tan^{-1} \frac{v_y}{v_x} = \tan^{-1} \frac{4.85}{7.20} \approx 34°$ □

Chapter 4

Force & Motion

4.1 force & motion: an introduction

in physics, a force appears when two bodies interact with one another

- you will encounter various types of forces in this course, some of which are
  - *weight*: the gravitational attraction acting on any object exerted by the earth
  - *tension*: a force in a string, a rope, a chain, etc. when it is being pulled
  - *normal contact*: when a body's surface is compressed, there reacts a normal contact force[17]
  - *friction*: a force that resists relative motion when two surfaces tend to slide over one another
  - *resistance*: also called drag force, this is experienced when a body travels through a medium
  - *upthrust*: an upward force acting on an object immersed in a fluid
  - *electric force*: an attractive or repulsive interaction between electrically charged objects

  detailed features of these forces follow later in the notes.
- a force can produce various effects to the object, the effect could be
  - an increase/decrease in speed
  - a change in the direction of motion
  - causing the object to rotate
  - a change in shape of the object

  in the next few chapters, we will be looking into each of these aspects
- when more than one force act on a body, it is useful to find their *resultant*, or the *net force*

resultant force, or **net force**, is a single force that has the same effect as all forces acting on an object combined

vector sum of all of the individual forces gives the resultant force

[17] Examples of normal contact force are support force that stops a desk from sinking into the ground, the impact on a football when you kick it, etc.

- in this chapter, we will study the dynamics of *point masses*

  point mass is an idealization that the object has a mass but does not take up any space

  position of an object treated as a point mass is specified with a geometric point in space

  this is a simplification when size, shape, rotation, or structure of object are not important

4.2 Newton's laws of motion

Newton's laws of motion[18] are three laws that form the basis of classical mechanics

they describe the relationships between motion of an object and forces acting on it

4.2.1 first law

> **Newton's first law** states that an object continues in its state of rest or uniform motion at constant velocity if there is no resultant force acting

- any object 'dislikes' any change to its state of motion, uniform motion tends to persist forever

  this tendency to resist changes in motion is called the **inertia**

  Newton's first law is also called *the law of inertia*
- if there is no change in state of motion, the object is said to be in **equilibrium**

  equilibrium could be either *static* (being at rest) or *dynamic* (steadily moving in a straight line)

  both cases require zero resultant force
- Newton's law of inertia is placed to establish frames of reference

  it is in an reference frame that notions of displacement, velocity and acceleration can be defined

  an **inertial reference frame** is one in which Newton's laws hold [19]

---

[18]These three laws were first addressed by *Isaac Newton* in his famous work *Mathematical Principles of Natural Philosophy*, or simply the *Principia*. The three-volume work was first published in 1687, and was soon recognised as one of the most important works in the history of science. Apart the from the three laws that laid the foundations for classical mechanics, the *Principia* also stated *the law of gravitation*, and accounted for planetary orbits and tides and other phenomena.

[19]Inertial frame is not unique. An observer moving at constant velocity with respect to an inertial observer is in a different inertial frame, since constant velocity added to a constant relative velocity is still a constant velocity. Two inertial observers would disagree on a body's velocity, but they would agree that the body maintains its velocity in absence of net force, so they will observe the same physics phenomena. This is

4.2.2 second law

if resultant force is non-zero, velocity of the object will change, i.e., force produces acceleration

Newton's second law states that the acceleration is proportional to the resultant force and inversely proportional to the mass of the object

- symbolically, we write $a \propto \dfrac{F_{\text{net}}}{m}$
  with consistent units of measure, this proportionality can be written as an exact equation:
$$a = \frac{F_{\text{net}}}{m} \quad \text{or} \quad \boxed{F_{\text{net}} = ma} \tag{4.1}$$
- SI unit of measurement for force F is **newton** (N)
  a force of one newton acting a body of 1 kg produces an acceleration of 1 m s^{-2}
- note that the force in the equation $F = ma$ is the resultant force
  to determine change in motion for a body, you should always ask what the resultant force is
- acceleration produced is always in same direction of the net force
- for same force, an object of greater mass has a smaller acceleration
  hence mass is a measure of the *inertia* of this object in response to a net force
  a definition for mass of an object from the point of view of Newton's laws can be stated as[20]:

mass is an intrinsic property of a body to resist any change in its state of motion

Example 4.1 A box of 6.0 kg is being pushed along a horizontal surface with a force of 30 N. The resistive force acting is 21 N. What is the acceleration of the box?

✍ $$F_{\text{net}} = F - f = ma \quad \Rightarrow \quad a = \frac{F - f}{m} = \frac{30 - 21}{6.0} = 1.5 \text{ m s}^{-2}$$ □

Example 4.2 A car of mass 800 kg is travelling at a speed of 20 m s^{-1}. The driver then operates the brake pedal so a braking force of 2000 N gradually brings the car to stop. (a) What is the deceleration for the car? (b) What is the stopping distance?

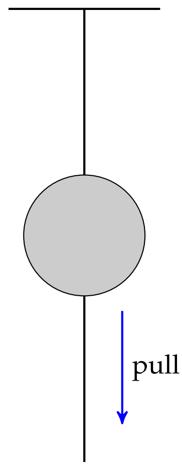
known as the *equivalence principle*.

[20]The concept of mass can be defined in many different ways. You might be familiar with the definition for mass as the amount of matter an object possesses. I personally think this definition is a bit vague and does not tell you anything new. Thinking of mass as a measure of inertia surely brings more insights. Mass also tells the strength at which an object interacts with other bodies through the gravitational attraction. As you will see later, from the view of Albert Einstein, it is also possible to think of mass as a form of energy.

✍ using Newton's second law and noticing braking force acts opposite to direction of motion:

$$F_{\text{net}} = ma \quad \Rightarrow \quad -2000 = 800 \times a \quad \Rightarrow \quad a = -2.5 \text{ m s}^{-2}$$
$$2as = v^2 - u^2 \quad \Rightarrow \quad s = \frac{v^2 - u^2}{2a} = \frac{0^2 - 20^2}{2 \times (-2.5)} = 80 \text{ m}$$ □

Example 4.3 A massive ball is suspended on a string. A second string is attached to the bottom of the ball. If one pulls the bottom string with a gradually increasing force, does the top string or the bottom string break first? What if the bottom string is jerked, which string breaks?



✍ when tension gradually increases, system is always in equilibrium

tensions in strings must have $T_{\text{top}} = T_{\text{bottom}} + mg$

top string suffers a greater force, so it breaks first

however, when bottom string is jerked, the ball tends to remain at rest due to its large mass, preventing sudden change to the tension in top string

so in this case bottom string is more likely to snap □

4.2.3 third law

every force is part of a pair of interactions between one body and another

when one body exerts a force on another, the second body also exerts a reaction on the first

> **Newton's third law**, also called the **action-reaction principle**, states that action and reaction are always equal in magnitude, opposite in directions and of the same type

Example 4.4 Suggest the action and reaction force in the following cases: (a) A man stands on a bathroom scale. (b) A helicopter hovers in air. (c) The earth orbits around the sun.

✍ (a) man exerts downward force on scale, scale exerts an upward reaction on man

(b) rotors of helicopter push air downwards, air exerts an upward force on helicopter

(c) sun pulls the earth through gravitational attraction, earth also attracts the sun in return □

4.2.4 force analysis & free-body diagrams

when doing mechanics problems, it is necessary to find all forces applied upon an object

to visualise all these forces, it is helpful to draw a **free-body diagram** (FBD)

an FBD shows a simplified version of the body with arrows indicating forces applied it is recommended to follow the routine stated below when solving a mechanics problem

(1) draw a FBD for the object in the problem

(2) resolve and find the resultant force with aid of the FBD

(3) apply Newton's laws to write down the equation of motion for the object

(4) solve the equation(s) to find acceleration

(5) use kinematic relations to deduce information about motion of the object

4.3 types of forces

4.3.1 weight

all objects exert attractive forces of gravity upon each other[21]

weight of a body is due to the gravitational pull from our planet – the earth

weight W of any object is proportional to its mass m: $\boxed{W = mg}$

g is strength of the gravitational field, or the gravitational acceleration constant

➢ at vicinity of earth's surface, gravitational field is almost uniform: $g \approx 9.81\ \mathrm{N\ kg^{-1}}$

but this value for g does not hold in a satellite orbit, on Mars, near a black hole, etc.

➢ the concept of weight is different from mass in many aspects

- weight is a force, so it is a vector (always acting downwards still makes a direction)
  mass is a scalar, it has magnitude only
- weight is measured in newtons, mass is measured in kilograms
- weight of object depends on its mass but also strength of gravitational field
  mass is an intrinsic property of object, so does not depend on force fields
  same object can have different weights on different planets, but its mass will be the same[22]

Example 4.5 An astronaut finds that he weighs 300 N on the surface of Mars, where the gravitational field strength is known to be 3.7 N kg^{-1}. Find his mass and hence estimate his weight if he returns to his home on the Earth.

✍ mass of astronaut: $m = \dfrac{W_M}{g_M} = \dfrac{300}{3.7} \approx 81.1\ \mathrm{kg}$

weight on earth: $W_E = mg_E = 81.1 \times 9.81 \approx 795\ \mathrm{N}$ □

[21] You will learn more about gravitational forces at A2 Level.

[22] Here we do not take into account the effects of *relativity*. A clever student who has learned Einstein's theories might suggest the mass of the same object increases with its velocity.

free fall

all things on the earth fall because of the force of gravity

if we ignore the restraints such as air resistance and upthrust force on a falling object, say the object is under the influence of gravity only, then the object is in a state called **free fall**

assuming the object is subject to gravity only, the resultant force is simply its weight

applying the Newton's second law, we have: $F_{\text{net}} = W \quad \Rightarrow \quad ma = mg$

so acceleration of the freely-falling object is: $\boxed{a = g}$[23]

➢ this shows **acceleration due to free fall** is simply equal to field strength g

so any object, regardless of its mass, has same acceleration due to free fall[24]

4.3.2 drag

when a body moves through air, water or any fluid, it experiences resistance called drag force

➢ factors that determine the value of fluid drag include

- relative speed of the object to the fluid ($v \nearrow \Rightarrow f \nearrow$)
- cross section of the object ($A \nearrow \Rightarrow f \nearrow$)
- shape of the object (streamlined shape has smaller drag)
- density of the fluid ($\rho \nearrow \Rightarrow f \nearrow$)

but what determines the drag force is a complicated issue[25]

[23]In the derivation, the mass terms cancel out. Rigorously speaking, these are two different masses. One is the measure of inertia, and the other is a measure of gravitational force. It is experimentally found that the inertia mass and the gravitational mass are equal. The fact that the two masses are equal has profound reasons. We have shown here acceleration of free fall equals gravitational field strength, but Albert Einstein's suggests that it is actually impossible to distinguish between a uniform acceleration and a uniform gravitational field. This idea lies at the heart of his *general theory of relativity*. Those who are interested in this topic are recommended to start from here and do some online researches.

[24]In §3.4.2 and §3.5, the statement that acceleration of free fall is constant in absence of air resistance was asserted without further explanation. Now you know why.

[25]There are a few empirical formula for drag force, each of which is accurate under certain conditions.

For an object moving through a fluid at low speeds (*laminar flow*, no turbulence occurs), the resistance it experiences is proportional to its speed: $f = bv$, where b is some constant which depends on fluid viscosity and the effective cross-sectional area of the object.

If objects are moving at relative high speeds through the fluid such that *turbulence* is produced behind the

➢ drag force always acts in a direction to oppose relative motion of object through fluid

free fall through air

let's consider an object falling through air from a very high tower

forces acting are weight and air resistance (shown in the free-body diagram)

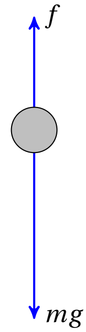
equation of motion for this falling object is:

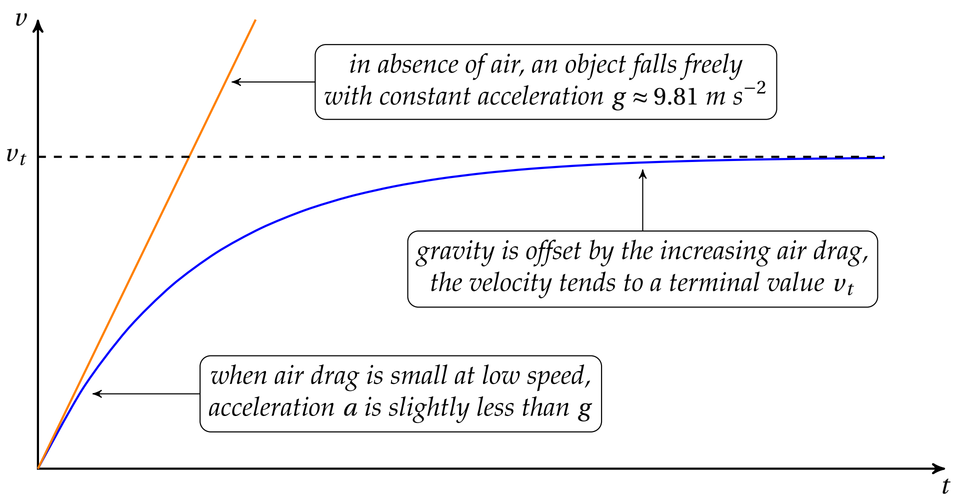
$$F_{\text{net}} = mg - f = ma$$

as v increases, air resistance f increases, so net force F_{net} decreases

this means acceleration a would decrease as object falls

i.e., speed will increase at a decreasing rate during the fall [26]





variation of velocity for a falling object through air

---

object, drag force is proportional to the speed squared: $f = \frac{1}{2}\rho C_D A v^2$, where ρ is the fluid's density, A is the cross-sectional area, C_D is a dimensionless quantity called the drag coefficient.

[26]The velocity-time relation can be obtained for some simple models. Suppose air resistance is proportional to speed of the falling body, i.e., $f = bv$, then the equation of motion reads: $F_{\text{net}} = m\frac{\mathrm{d}v}{\mathrm{d}t} = mg - bv$, where acceleration is written explicitly as the rate of change in velocity. With the initial conditions $v = 0$ at $t = 0$, we can solve this differential equation to obtain the speed of this falling object at any given time t:

$$\mathrm{d}t = \frac{\mathrm{d}v}{g - \frac{b}{m}v} \quad \Rightarrow \quad \int_0^t \mathrm{d}t = \int_0^v \frac{\mathrm{d}v}{g - \frac{b}{m}v} \quad \Rightarrow \quad t = -\frac{m}{b}\ln\left(g - \frac{b}{m}v\right)\Big|_0^v = -\frac{m}{b}\ln\left(1 - \frac{bv}{mg}\right)$$

Rearrange the terms, we find: $v(t) = \frac{mg}{b}\left(1 - \mathrm{e}^{-\frac{bt}{m}}\right)$

➢ at low speeds, air resistance is negligible, so $F_{\text{net}} = ma \approx mg$

acceleration of object at start of the fall is similar to g

but as v increases, acceleration decreases so a becomes less than g

➢ after sufficient long time, acceleration gradually decreases to zero

velocity gradually increases and tends to a maximum value

at this stage, equilibrium is restored: $f = mg$, object no longer accelerates

this constant final velocity is known as the **terminal velocity**

Example 4.6 An object of 5.0 kg falls through the atmosphere from a very high altitude. After some time, it falls at a constant speed of 70 m s^{-1}. Assume there is no significant change in gravitational field during the fall and the air resistance is proportional to speed: $f = bv$. (a) Find the value of the coefficient k. (b) Find the acceleration of the object when it is falling at 30 m s^{-1}.

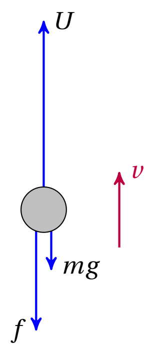
✍ equilibrium between weight and air drag when falling at terminal speed, so

$$mg = bv_t \quad \Rightarrow \quad b = \frac{mg}{v_t} = \frac{5.0 \times 9.81}{70} \approx 0.70 \text{ kg s}^{-1}$$

at any instant, equation of motion is: $F_{\text{net}} = ma = mg - bv$

at 30 m s^{-1}, acceleration is: $a = \dfrac{mg - bv}{m} = \dfrac{5.0 \times 9.81 - 0.70 \times 30}{5.0} \approx 5.6 \text{ m s}^{-2}$ □

bubble rising in a liquid



let's now consider bubbles formed at the bottom of a soda water

forces acting on bubble are weight, water resistance and upthrust

equation of motion for the rising bubble is:

$$F_{\text{net}} = U - mg - f = ma$$

as bubble moves faster, f increases, then F_{net} decreases

so acceleration a would gradually decrease to zero as bubble rises

speed of bubble increases and reaches a maximum value

at terminal speed, $a \to 0$, one has: $U = f + mg$

4.3.3 normal contact

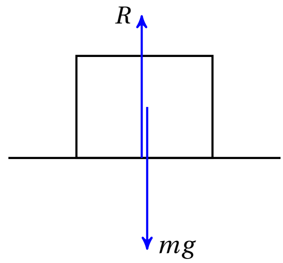
when two objects are in contact, the interaction between them is called the *contact force*

normal contact force is the component of contact force perpendicular to the contacting surface

➢ by definition, normal contact is always at right angle to surface of contact

➢ origin of normal contact is the *electrostatic interaction* between atoms when two objects are pressed against each other, surface atoms get close

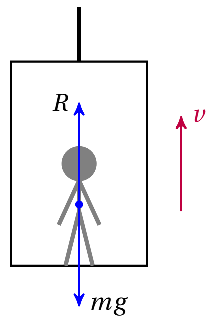
electrostatic repulsion between electron clouds of the atoms prevent them from penetrating through one another



Example 4.7 A box of mass $m = 4.0$ kg is resting on a horizontal ground. What is the normal contact force acting?

✍ equilibrium between weight and normal contact, so

$$R - W = 0 \quad \Rightarrow \quad R = mg = 4.0 \times 9.81 \approx 39.2 \text{ N} \quad \square$$



Example 4.8 A man of 80 kg stands in a lift. Find his apparent weight, i.e., the contact force, when the lift is (a) moving upwards at steady speed of 2.0 m s^{-1}, (b) accelerating upwards at 2.0 m s^{-2}, (c) moving upwards but slowing down at a deceleration of 1.5 m s^{-2}.

✍ forces acting on man are weight and normal contact

for either case, equation of motion for the man reads:

$$F_{\text{net}} = ma = R - mg$$

so normal contact force: $R = mg + ma$

when rising at steady speed, man is in equilibrium ($a = 0$), so: $R = mg = 80 \times 9.81 \approx 785$ N

when accelerating upwards ($a = +2.0$ m s^{-2}): $R = 80 \times 9.81 + 80 \times 2.0 \approx 945$ N

when decelerating upwards ($a = -1.5$ m s^{-2}): $R = 80 \times 9.81 + 80 \times (-1.5) \approx 665$ N $\square$

Example 4.9 A sleigh of mass 15 kg lies at rest on an icy ground. The surface is frictionless. A force P of 75 N is applied to the sleigh. Find the normal contact force and the acceleration of the sleigh if P is acting (a) horizontally, (b) at an angle α to the horizontal where $\tan\alpha = \frac{3}{4}$.

✍ free-body diagrams for both cases are shown

for both cases, no net force acts in vertical direction

net force in horizontal direction provides acceleration

when P acts horizontally:

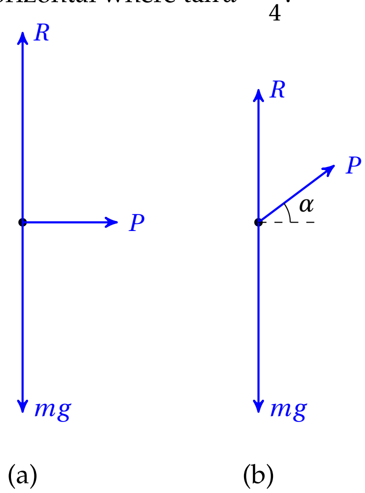
$$R = mg = 15 \times 9.81 \approx 147 \text{ N}$$

$$P = ma \quad \Rightarrow \quad a = \frac{P}{m} = \frac{75}{15} = 5.0 \text{ m s}^{-2}$$

when P acts at angle α:

$$R + P\sin\alpha = mg \quad \Rightarrow \quad R = 15 \times 9.81 - 75 \times \frac{4}{5} \approx 87 \text{ N}$$

$$P\cos\alpha = ma \quad \Rightarrow \quad a = \frac{P\cos\alpha}{m} = \frac{75 \times \frac{3}{5}}{15} = 3.0 \text{ m s}^{-2} \quad \square$$



4.3.4 friction

friction is the component of contact force that is parallel to contact surfaces

when there is potential or actual sliding between surfaces, frictional force come into action

– for surfaces *tend* to move relative to each other, **static friction** acts to oppose this tendency

– if surfaces are already sliding over one another, then **kinetic friction** opposes this motion

➢ static friction f_S is self-adjusting

an object placed on a rough surface can stay at rest when acted by a small external force F

it can do so because f_S equalises external force to maintain equilibrium

if no external force acts, then $f_S = 0$

➢ there exists a maximum limiting friction f_{lim}

when external force $F < f_{\text{lim}}$, there is sufficient f_S to prevent object from sliding

when $F = f_{\text{lim}}$, object is on the verge of sliding

when $F > f_{\text{lim}}$, object start to move and static friction becomes kinetic friction f_K

➢ factors that determine frictional forces are

– nature of contacting surfaces (for both f_S and f_K)

– normal reaction R (for f_K)

these dependences are usually expressed by a mathematical equation $\boxed{f_K \approx f_{\text{lim}} = \mu R}$

μ is the *coefficient of friction* whose value depends on the nature of the two surfaces[27]

➢ friction, on microscopic level, is an *electromagnetic force* in nature

when two surfaces are in contact, irregularities on the surface touch each other

surface atoms come very close and bonds are formed through electrostatic force

in some sense, surface atoms get *cold welded* to each other

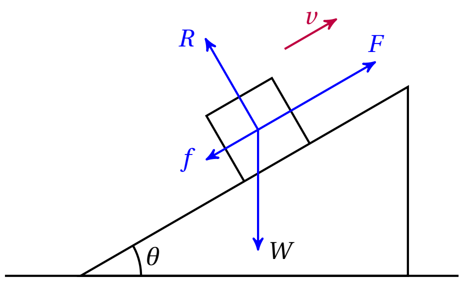
when surfaces try to move relative to each other, this electrostatic weld is origin of friction

4.4 inclined slope

inclined slope is probably the entry ticket into the business of mechanics

this notorious problem is found in any physics textbook and any exam paper on mechanics

[27]The idea of limiting friction is not required in the AS Physics syllabus, but it is required in *Mechanics 1* of the A-Level Mathematics course.



the problem is about a mass m placed on a plane inclined at angle θ to the horizontal

the mass could sit at rest on, slide down, or get pulled/pushed up the plane

motion of the mass could be affected by weight, friction, normal contact, or other forces

➢ forces can be resolved in directions parallel and perpendicular to the slope

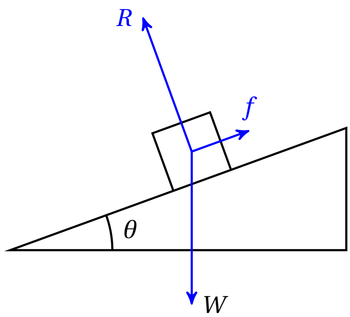
resolving along the slope leads to the equation of motion from which acceleration is found

➢ it is almost inevitable to break weight into two components[28]

– component of weight parallel down the slope is: $W_{/\!/} = mg\sin\theta$

– component of weight perpendicular to slope is: $W_\perp = mg\cos\theta$

Example 4.10 A block of mass m stays at rest on an inclined plane. The plane makes an angle θ with the horizontal. Find the normal contact force R and the frictional force f acting on the block.

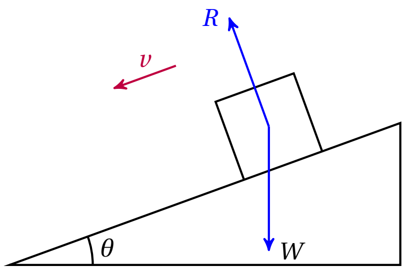


✍ block in equilibrium, so $F_{\text{net}} = 0$ in any direction

parallel to slope: $f = W_{/\!/} \quad\Rightarrow\quad f = mg\sin\theta$

normal to slope: $R = W_\perp \quad\Rightarrow\quad R = mg\cos\theta$ □

Example 4.11 A block of mass m slides down a *smooth* slope. The angle of the slope to the horizontal is θ. Find the acceleration of the block.



✍ only force acting along the slope is component of weight down the slope, so:

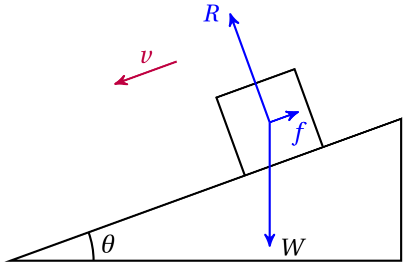
$$F_{\text{net}} = ma = mg\sin\theta \quad\Rightarrow\quad a = g\sin\theta$$

as $\theta \to 0$, $a \to 0$, this shows if plane becomes horizontal, the block simply stays put

as $\theta \to 90^\circ$, slope becomes vertical, block would undergo free fall, so naturally $a \to g$ □

[28] We have already done that in Example 1.7.

Example 4.12 A block of mass 2.0 kg slides down a rough slope from rest. The slope is inclined at angle $\theta = 20^\circ$ to the horizontal, and the block experiences a constant friction of 5.0 N. (a) What is the block's acceleration? (b) What is the distance travelled in 2.5 seconds?



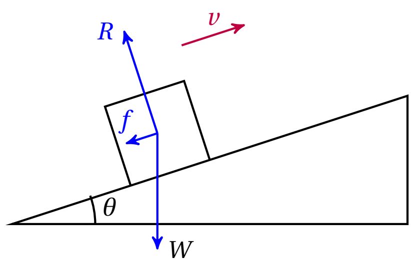
✍ resolving along slope:

$$F_{\text{net}} = mg\sin\theta - f = ma$$

$$a = \frac{2.0 \times 9.81 \times \sin 20^\circ - 5.0}{2.0} \approx 0.855 \text{ m s}^{-2}$$

distance travelled: $s = \cancelto{0}{ut} + \frac{1}{2}at^2 = \frac{1}{2} \times 0.855 \times 2.5^2 \approx 2.67 \text{ m}$ □

Example 4.13 A block of mass 3.0 kg is travelling up an inclined slope at an initial speed of 2.8 m s^{-1}. The slope makes an angle of 18° with the horizontal. A constant friction of 7.5 N acts on the block. (a) What is the block's deceleration? (b) How far does the block travel along the slope before its speed decreases to zero? (c) Suggest whether the block could stay on the slope.



✍ resolving along slope (take direction of initial velocity as positive direction):

$$F_{\text{net}} = -mg\sin\theta - f = ma \quad \Rightarrow \quad a = \frac{-mg\sin\theta - f}{m} = \frac{-3.0 \times 9.81 \times \sin 18^\circ - 7.5}{3.0} \approx -5.53 \text{ m s}^{-2}$$

$$v^2 - u^2 = 2as \quad \Rightarrow \quad s = \frac{v^2 - u^2}{2a} = \frac{0^2 - 2.8^2}{2 \times (-5.5.3)} \approx 0.709 \text{ m}$$

note that component of weight down the slope is: $W_{/\!/} = mg\sin\theta = 3.0 \times 9.81 \times 18^\circ \approx 9.1$ N

$W_{/\!/} > f$, so friction is not enough to prevent block from sliding back down the slope □

4.5 many-body problems

the problems we have been dealing with so far only involve one body

a mechanical system could consist of several objects that mutually interact

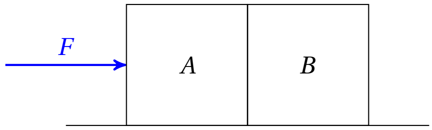
➢ one can take each individual and look into the *internal* forces between the objects of interest

for any force acting between objects *within* system, there is an equal but opposite reaction force

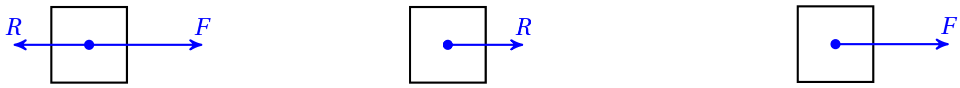
➢ the system can also be treated as a whole

we can analyse *net external force* acting on entire system and work out combined acceleration

Example 4.14 Two boxes A and B are placed on a smooth surface. They are accelerated together by a horizontal force F as shown. Find the acceleration and the contact force between them.



✍ free-body diagrams for A, B, and entire system are given below



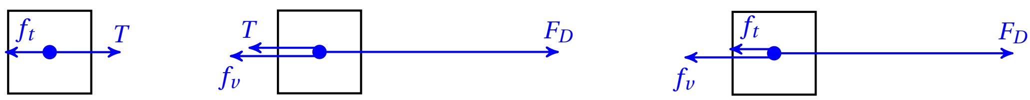
A B A and B combined

equations of motion can be written down for each free-body diagram and solved[29]

$$\begin{cases} \text{for } A: & F - R = M_A a \\ \text{for } B: & R = M_B a \\ \text{for system:} & F = (M_A + M_B)a \end{cases} \Rightarrow \begin{cases} a = \dfrac{F}{M_A + M_B} \\ R = \dfrac{M_B}{M_A + M_B} F \end{cases}$$ □

Example 4.15 A vehicle of mass 1500 kg is towing a trailer of mass 500 kg by a light inextensible tow-bar. The engine of the vehicle exerts a driving force of 9600 N, and the tractor and the trailer experience resistances of 3600 N and 1800 N respectively. Find the acceleration of the vehicle and the tension in the tow-bar.

✍ free-body diagrams for trailer, vehicle and entire system are given below



trailer vehicle system

equations of motion can be written down for each free-body diagram:

$$\begin{cases} \text{for trailer:} & T - f_t = M_t a \\ \text{for tractor:} & F_D - f_v - T = M_v a \\ \text{for system:} & F_D - f_v - f_t = (M_v + M_t)a \end{cases} \Rightarrow \begin{cases} T - 1800 = 500a \\ 9600 - 3600 - T = 1500a \\ 9600 - 3600 - 1800 = (1500 + 500)a \end{cases}$$

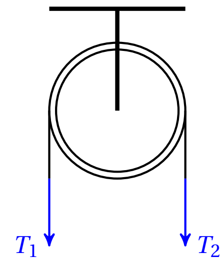
solving simultaneous equations[30], we find

$$a = 2.1 \text{ m s}^{-2}, \quad \text{and} \quad T = 2850 \text{ N}$$ □

[29] In fact, only two of the three equations are independent. You can easily check that adding the equation for A to that for B would produce the equation for the system. To solve the two unknowns for this problem, any two of the three equations shall do the job.

[30] Again, only two of the three equations are independent. You can freely choose your favourite two.

pulleys



a *pulley* is basically a wheel that carries a string/rope/cable

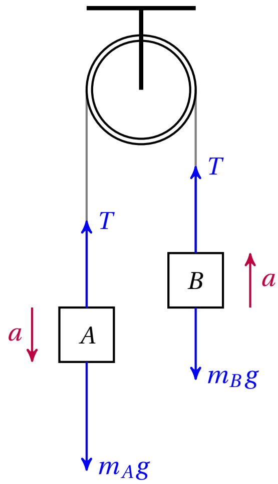
in this section, we only consider pulleys whose axis of rotation is fixed

such pulleys can be used to change direction of tension in a taut string

we also assume pulleys to be *ideal*: they have no mass and no friction

for an ideal pulley, tensions on both sides are equal: $T_1 = T_2$

Example 4.16 Two blocks of mass m_A and m_B ($m_A > m_B$) are joined together by a light inextensible string. The string passes over a smooth pulley as shown. The two blocks are suddenly released from rest. Find the acceleration of each block and the tension in the string.



✍ apply Newton's second law to each block:

$$\begin{cases} \text{for } A: & m_A g - T = m_A a \\ \text{for } B: & T - m_B g = m_B a \end{cases}$$

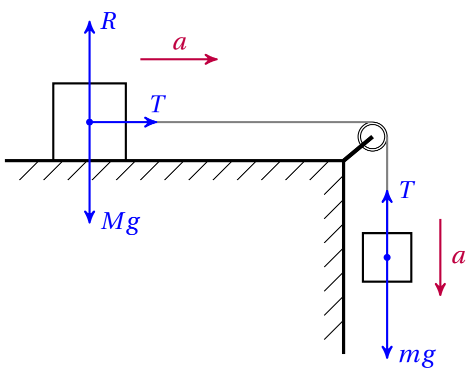
adding the two, one obtains equation of motion for whole system:

$$m_A g - m_B g = (m_A + m_B) a$$

solving these equations, we find

$$a = \frac{m_A - m_B}{m_A + m_B} g \qquad T = \frac{2 m_A m_B g}{m_A + m_B}$$ □

Example 4.17 A mass $M = 4.0$ kg is attached to a block of mass $m = 2.0$ kg through a light string which passes over a frictionless pulley as shown. When both masses are released, find the acceleration and the tension in the string.



✍ apply Newton's second law to each mass:

$$\begin{cases} \text{for } M: & T = Ma \\ \text{for } m: & mg - T = ma \end{cases}$$

adding the two equations, we have:

$$mg = (M + m) a$$

so acceleration is:

$$a = \frac{mg}{M+m} = \frac{2.0 \times 9.81}{4.0 + 2.0} = 3.27 \text{ m s}^{-2}$$

tension in string: $T = Ma = 4.0 \times 3.27 \approx 13.1$ N □

CHAPTER 5

Mechanical Equilibrium

in this chapter, we will study the mechanical equilibrium of objects

for *point objects*, zero resultant force suffices for equilibrium

but for *rigid bodies*, force many produce *turning effects*

hence rigid bodies must satisfy another condition to stay in equilibrium

this brings forward the notion of moment of a force and the principle of moments

5.1 moment of force

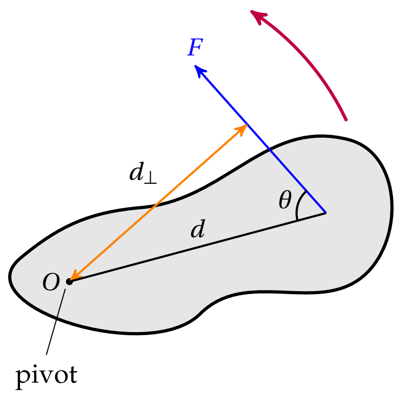
torque, or **moment** of a force, is defined as the product of the force and the perpendicular distance from the pivot to the line of action: $\boxed{\tau = F d_\perp}$

- unit of torque/moment: $[\tau] = \mathrm{N\,m}$
- perpendicular distance from pivot to the line of action is also called the **lever arm**

  this is the shortest distance between the force applied and axis of rotation

  in the diagram, lever arm $d_\perp = d \sin\theta$

  so moment of this force is: $\tau = F d \sin\theta$



- moment is a vector quantity [31] [32]

  it can act in *clockwise* or *anti-clockwise* direction

[31]Using vector notation, moment of a force can be defined as a *cross product*: $\vec{\tau} = \vec{r} \times \vec{F}$, where $\vec{r}$ is the position vector from the pivot to the point at which the force is applied.

[32]Rigorously speaking, moment is a *pseudovector*, which means that it does not transform quite like a normal vector although it does have a direction. In particular, if an object acted by a force is reflected across a plane, the moment of this force would not be reflected. Instead, it would be reflected and *reversed*.

➢ moment of a force produces turning effects[33]

if there exists a non-zero moment, the object will start to rotate clockwise or anti-clockwise

➢ note that moment of a force depends on choice of pivot

moment of the same force with respect to different points can be very different

5.1.1 torque of couple

let's take a pair of equal but opposite forces acting at different positions on the same object

the two forces produce torques in the same direction

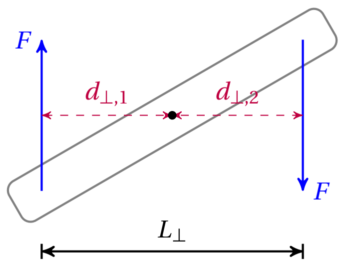
the combined effect is called the torque of a couple

resultant torque due to the couple is

$$\tau = Fd_{\perp,1} + Fd_{\perp,2} = 2F(d_{\perp,1} + d_{\perp,2})$$

$d_{\perp,1} + d_{\perp,2}$ is perpendicular distance $L_\perp$ between the pair, so

$$\boxed{\tau = FL_\perp}$$



torque of a couple can be therefore defined as the product of one force in the couple and the perpendicular distance between the pair

➢ a pair of equal but opposite forces give zero net force

but they can produce rotational effects, so no net force does not necessarily mean equilibrium

➢ torque of couple does not depend on choice of pivot

for same force pair, resultant moment is constant about any point

5.1.2 moment of weight

recall that weight is a force of gravity which is actually experienced by *all* parts of the object

when dealing with moment of weight, we need to sum up torques on each part of this object

this brings a problem since the lever arms can be all different

fortunately, this calculation can be simplified using the idea of centre of gravity

centre of gravity is a point at which the entire weight of an object is considered to act

[33]Moment is like the rotational counterpart of a force: force changes the state of translational motion, moment changes the state of rotation.

➢ there is a similar concept called the centre of mass

centre of mass is the average position of all the mass that makes up the object

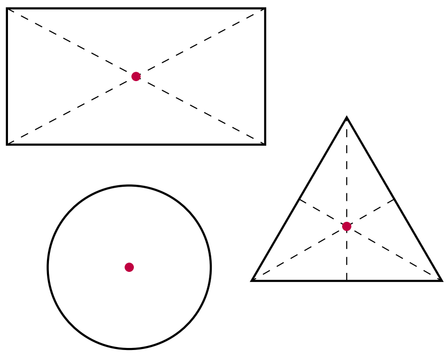
near the surface of the earth, mass and weight are directly proportional to each other

so centre of mass is interchangeable with centre of gravity if we stay on earth

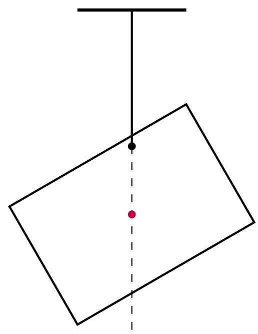
➢ for a regularly-shaped uniform object, the centre of gravity/mass is its geometrical centre

➢ if an object is hung freely, centre of gravity/mass is vertically below the point of suspension

otherwise weight would produce a non-zero torque about the point of suspension, causing the object to rotate until torque becomes zero



centre of mass of uniform lamina is at the geometrical centre



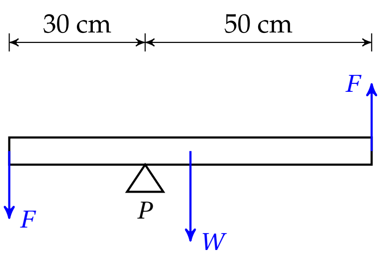
centre of mass is vertically below the point of suspension

➢ to find the centre of gravity/mass, the object of interest is suspended from several positions

each time we draw a vertical *plumb-line* through the point of suspension

centre of mass/gravity lies where the lines intersect

Example 5.1 The diagram shows a uniform beam of weight $W = 20$ N and length 80 cm pivoted at point P. P is 30 cm from one end. Two equal but opposite forces of magnitude $F = 12$ N are acting at the two ends of the beam as shown. What is the resultant moment about point P?



✍ moment of weight: $\tau_w = Wd_w = 20 \times \left(0.50 - \frac{1}{2} \times 0.80\right) = 2.0$ N m (clockwise)

torque of couple: $\tau_c = FL = 12 \times 0.80 = 9.6$ N m (anti-clockwise)

resultant moment: $\tau_{\text{net}} = \tau_c - \tau_W = 9.6 - 2.0 = 7.6$ N m (anti-clockwise) □

5.2 mechanical equilibrium

5.2.1 principle of moments

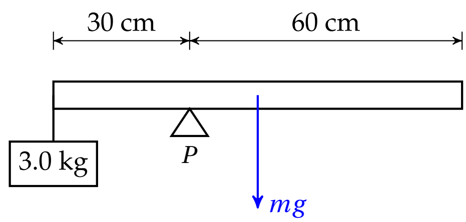
if there is no turning effect for an object, the total moment of all forces must vanish

> for a rigid body in equilibrium, sum of all clockwise moments must be equal to the sum of anti-clockwise moments *about any point*, this is called the **principle of moments**

➢ an object in equilibrium has no turning effect about any point

so resultant moment is zero about any point [34]

Example 5.2 A uniform rod of length 90 cm is pivoted 30 cm from one end. It is balanced with a 3.0 kg load. Find the mass of the rod.

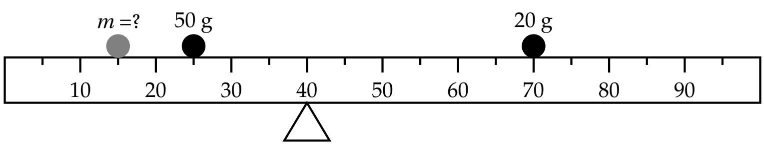


✍ take moments about P:

$$3.0 \times 9.81 \times 0.30 = m \times 9.81 \times \left(0.60 - \frac{1}{2} \times 0.90\right)$$

so we find mass of rod: $m = 1.5$ kg □

Example 5.3 A student balances a metre rule of mass 120 g supported on a fulcrum at the 40 cm mark. She then places a 20 g mass on the 70 cm mark and a 50 g mass on the 25 cm mark as shown. To balance the rule, what mass should she place on the 15 cm mark?



[34] As long as there is no resultant force, then zero resultant moment about any particular point would imply zero resultant moment about any point.

Mathematically, let's take a collection of forces $\vec{F_1}, \vec{F_2}, \cdots, \vec{F_n}$ acting at positions $\vec{r_1}, \vec{r_2}, \cdots, \vec{r_n}$ on an object with respect to some fixed point O. Suppose their resultant moment vanishes, i.e., $\sum \vec{\tau_i} \equiv \sum \vec{r_i} \times \vec{F_i} = 0$, and also their resultant force vanishes, i.e., $\sum \vec{F_i} = 0$. If we focus on a different point O' with a relative displacement $\vec{R}$ to point O', then taking moments about O', we will have:

$$\sum \vec{\tau_i'} = \sum \vec{r_i'} \times \vec{F_i} = \sum (\vec{r_i} + \vec{R}) \times \vec{F_i} = \sum \vec{r_i} \times \vec{F_i} + \vec{R} \times \sum \vec{F_i} = 0 + 0 = 0$$

which shows zero resultant moment about one point together with zero resultant force guarantee resultant moment must be zero about any point in space.

✍ take moments about the support:

$$mg\times(40-15)+0.050g\times(40-25)=0.12g\times(50-40)+0.020g\times(70-40)\quad\Rightarrow\quad m=42\text{ g}\qquad\square$$

Example 5.4 A cylinder of weight 100 N and diameter 50 cm rests against point P of a curb of height 10 cm. What is the minimum force required to cause the cylinder to roll to the left?

✍ to roll the cylinder, force applied must produce a torque no less than that of weight

force is minimum if lever arm is greatest

take moments about P (see diagram):

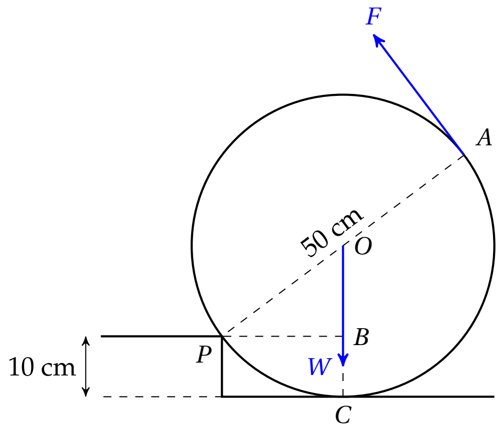
$$F_{\min}\times|PA|=W\times|PB|$$

note that $|PB|=\sqrt{|OP|^2-|OB|^2}$, so

$$|PB|=\sqrt{0.25^2-(0.25-0.10)^2}=0.20\text{ m}$$

plug back into the equation above:

$$F_{\min}\times 0.50=100\times 0.20\quad\Rightarrow\quad F_{\min}=40\text{ N}\quad\square$$



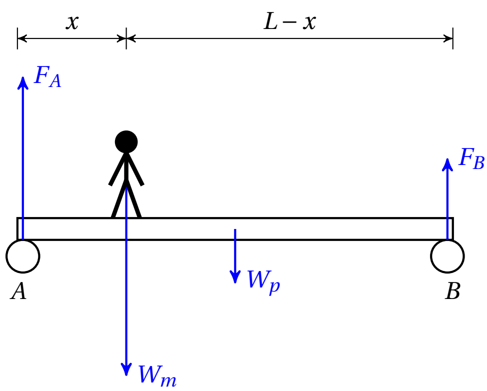
5.2.2 mechanical equilibrium

combining Newton's first law and principle of moments, we have the following statement:

> for any mechanical system in equilibrium, two conditions must be satisfied:
>
> – resultant force is zero in any direction: $\sum F=0$
>
> – resultant moment is zero about any point: $\sum\tau=0$

these two conditions allow for many possible equations that can be written down

Example 5.5 A uniform plank of weight 100 N and length $L=6.0$ m rests horizontally on two supports A and B. A man of weight 800 N stands a distance of $x=1.5$ m from end A. Determine the forces acting at the two supports.



✍ take moments about A: $W_m x+W_p\cdot\dfrac{L}{2}=F_B L$

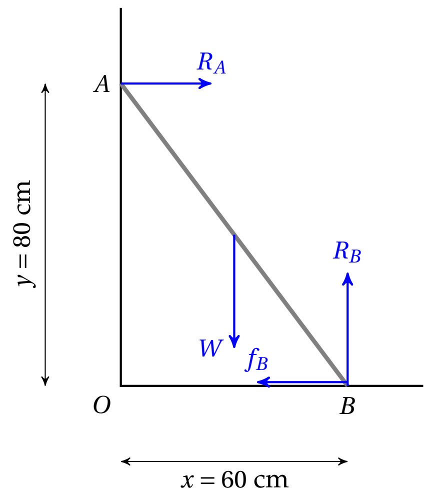
$800\times1.5+100\times3.0=F_B\times6.0\Rightarrow F_B=250\text{ N}$

take moments about B: $W_m(L-x)+W_p\dfrac{L}{2}=F_A L$

$800\times4.5+100\times3.0=F_A\times6.0\Rightarrow F_B=650\text{ N}$

one can check that: $F_A+F_B=W_m+W_p$, there must be no resultant force in vertical direction $\square$

Example 5.6 A uniform ladder of weight 120 N rests on a rough ground against a smooth wall as shown. The dimensions are labelled on the diagram. (a) What is the contact force acting at B? (b) What is the contact force acting at A? (c) What is the frictional force at B?



✍ free-body diagram is drawn as shown

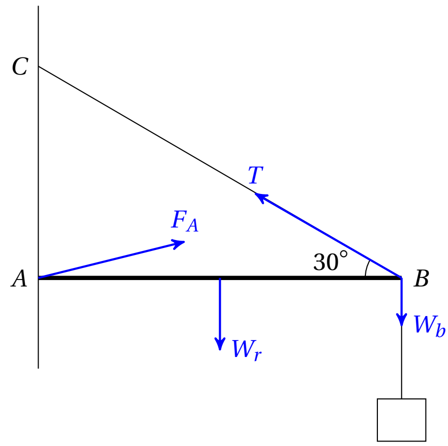
resolve vertically: $R_B = W \quad \Rightarrow \quad R_B = 120$ N

take moments about B: $R_A y = W\dfrac{x}{2}$

$$R_A = \frac{120 \times 0.30}{0.80} = 45 \text{ N}$$

resolve horizontally: $f_B = R_A \quad \Rightarrow \quad f_B = 45$ N □

Example 5.7 The diagram shows a uniform rod AB of weight 60 N being held horizontally to a vertical wall by means of a light string. The string is attached to the rod at B, where a basket of weight 40 N is suspended. The other end of the string is fixed on the wall at C. The angle between the string and the rod is 30°. (a) Find the tension in the string. (b) Find the force acting on the rod at point A.



✍ take moments about A: $TL\sin\theta = W_b L + W_r \dfrac{1}{2}L$

$$T\sin 30° = 40 + 60 \times \frac{1}{2} \quad \Rightarrow \quad T = 140 \text{ N}$$

resolve horizontally: $F_{A,x} = T\cos\theta \quad \Rightarrow \quad F_{A_x} = 140\cos 30° \approx 121$ N

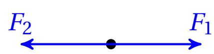
resolve vertically: $F_{A,y} + T\sin\theta = W_r + W_b \quad \Rightarrow \quad F_{A,y} = 60 + 40 - 140\sin 30° = 30$ N

force at A: $F_A = \sqrt{F_{A,x}^2 + F_{A,y}^2} \quad \Rightarrow \quad F_A = \sqrt{121^2 + 30^2} \approx 125$ N □

5.3 two forces in equilibrium

the problem of two balanced forces is trivial

suppose two forces F_1 and F_2 are acting on an object in equilibrium



- to have zero resultant force, F_1 and F_2 must be equal but opposite
- to have zero resultant moment, F_1 and F_2 must act along same line

  otherwise they produce torque of couple that causes turning effects

5.4 three forces in equilibrium

5.4.1 force triangle

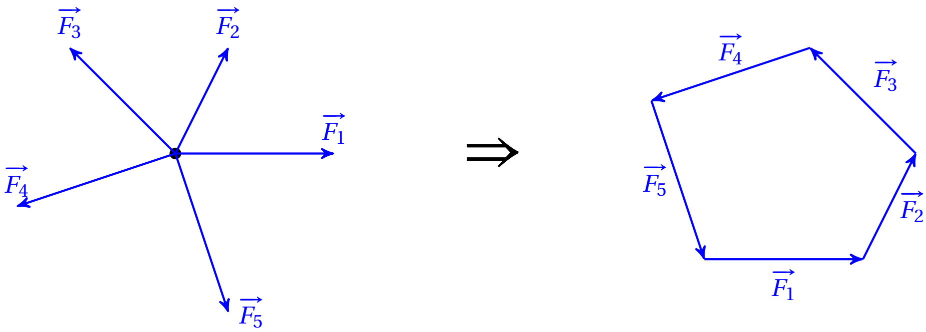
when there are more than two forces, situation becomes more complicated

one can use *vector diagram* to solve the problem

suppose a set of forces $\vec{F_1}, \vec{F_2}, \cdots, \vec{F_n}$ are in equilibrium

no resultant force requires $\vec{F_1} + \vec{F_2} + \cdots + \vec{F_n} = 0$

recall that resultant force is vector sum of all forces acting, and now this sum has to vanish, so if the force vectors are connected head to tail, they should form a closed n-polygon

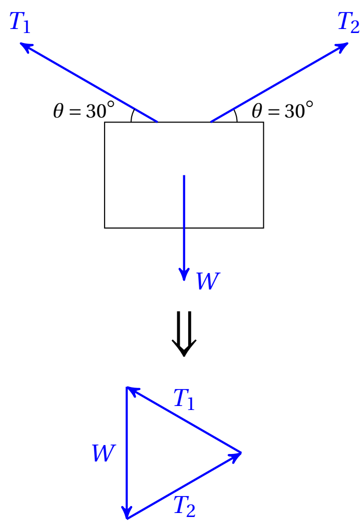


an n-polygon formed by a set of n balanced forces

in the case of three balanced forces, net force is zero means they should form a **force triangle**

unknown forces can then be solved by cracking a geometric problem

Example 5.8 A painting of weight $W = 20$ N is supported by two strings as shown. Both strings form an angle $\theta = 30^\circ$ to the horizontal. Find the tension in the strings.



✍ by resolving forces, we have:

$$\begin{cases} T_1 \cos\theta = T_2 \cos\theta \\ T_1 \sin\theta + T_2 \sin\theta = W \end{cases}$$

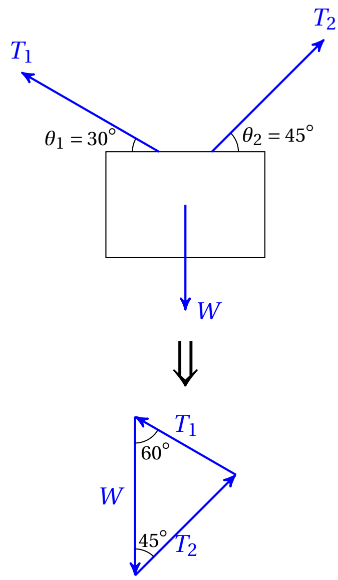
we solve the equations to obtain:

$$T_1 = T_2 = \frac{W}{2\sin\theta} = \frac{20}{2\sin 30^\circ} = 20 \text{ N}$$

alternatively, we can construct the force triangle as shown

T_1, T_2 and W form an equilateral triangle, so

$$T_1 = T_2 = W = 20 \text{ N}$$ □



Example 5.9 The same painting of weight $W = 20\,\mathrm{N}$ is supported by two strings at different angles $\theta_1 = 30^\circ$ and $\theta_2 = 45^\circ$ as shown. Find the forces in the two strings.

✍ by resolving forces, we have:

$$\begin{cases} T_1 \cos\theta_1 = T_2 \cos\theta_2 \\ T_1 \sin\theta_1 + T_2 \sin\theta_2 = W \end{cases} \Rightarrow \begin{cases} \frac{\sqrt{3}}{2} T_1 = \frac{\sqrt{2}}{2} T_2 \\ \frac{1}{2} T_1 + \frac{\sqrt{2}}{2} T_2 = 20 \end{cases}$$

solving this, we find: $T_1 \approx 14.6\,\mathrm{N}$, $T_2 \approx 17.9\,\mathrm{N}$

alternatively, we construct the force triangle as shown

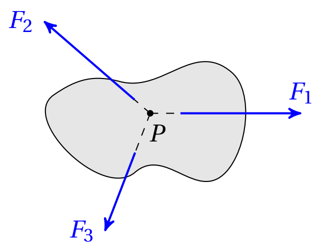
T_1 and T_2 are related to W by the the law of sine:

$$\frac{W}{\sin 75^\circ} = \frac{T_1}{\sin 45^\circ} = \frac{T_2}{\sin 60^\circ}$$

from this we get the same result: $T_1 \approx 14.6\,\mathrm{N}$, $T_2 \approx 17.9\,\mathrm{N}$ □

5.4.2 concurrent forces

for three forces in equilibrium, they must produce zero resultant moment about any point



suppose lines of action of F_1 and F_2 meet at point P

moment of F_1 and moment of F_2 about P are both zero

to produce zero resultant moment about P, then moment of F_3 about P must vanish

this suggest line of action of F_3 must pass through point P

therefore lines of action of F_1, F_2 and F_3 must pass through the same point[35]

such three forces are said to be *concurrent*

summary for three forces in equilibrium

for three forces in equilibrium, we can now conclude:

- the three force vectors must be able to form a force triangle

  this is a consequence of zero resultant force
- the lines of action for the three forces must pass through same point

  this is a consequence of zero resultant moment

[35] In the case of three parallel forces in equilibrium, we can introduce the notion of an ideal point at infinity, so that parallel lines could meet at that point.

CHAPTER 6

Momentum

6.1 momentum & momentum conservation

6.1.1 momentum

momentum of an object is defined as the product of its mass and its velocity: $\boxed{p = mv}$

- unit of momentum: $[p] = \mathrm{kg\ m\ s^{-1}} = \mathrm{N\ s}$
- momentum is a vector quantity

  momentum is in same direction as the object's velocity

  to find change in momentum of a body, or to find sum of the momenta[36] for a system of several objects, one has to keep track of directions
- momentum is like inertia in motion

  inertia, the property that object resists change in motion, is incorporated in Newton's first law

  we will see very soon that momentum is closely related to Newton's second law

6.1.2 relation to force

suppose a constant net force F is applied on a body, we can write:

$$F = ma = m\frac{\Delta v}{\Delta t} = \frac{m\Delta v}{\Delta t} = \frac{\Delta p}{\Delta t}$$

where we have used Newton's second law and defining equation for momentum

from this derivation, we can give a formal definition for the force

force is defined as the rate of change in momentum: $\boxed{F = \dfrac{\Delta p}{\Delta t}}$

[36]Momenta is the plural form of momentum.

we can also restate Newton's second law in terms of momentum:

> **Newton's second law** states that resultant force acting on an object equals the rate of change in the object's momentum

- information about force can be extracted from momentum-time graphs

  gradient of p-t graph equals the resultant force acting

- information about *change* in momentum can be deduce from force-time graphs

  area under F-t graph equals the change in object's momentum

Example 6.1 A ball of 120 g strikes a wall at right angle with a speed of 10 m s^{-1}. It rebounds with the same speed. If the time of impact is 25 ms, find the average force exerted on the ball.

✍ change in momentum: $\Delta p = mv - mu = 0.12 \times 10 - 0.12 \times (-10) = 2.4$ kg m s^{-1}

average force: $F = \dfrac{\Delta p}{\Delta t} = \dfrac{2.4}{2.5 \times 10^{-3}} = 96$ N □

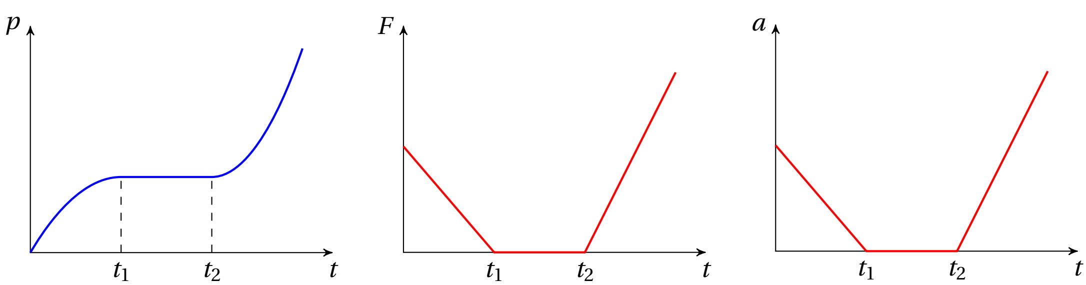
Example 6.2 Water is pumped through a hose-pipe. A man is holding the hose-pipe horizontally and water emerges from the hose-pipe with a speed of 16 m s^{-1} at a rate of 45 kg per minute. Find the force required from this man to hold steady the hose-pipe.

✍
$$F = \frac{\Delta p}{\Delta t} = \frac{\Delta(mv)}{\Delta t} = \frac{\Delta m(v-0)}{\Delta t} = \frac{45 \times 16}{60} \quad \Rightarrow \quad F = 12 \text{ N}$$
□

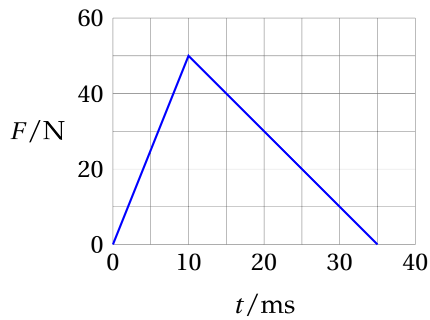
Example 6.3 A strong wind of speed 30 m s^{-1} blows against a wall of area 10 m^2 at right angles. The density of the air is 1.2 kg m^{-3}. Assume air speed reduces to zero when it hits the wall. What is the approximate force exerted by the air on the wall?

✍
$$F = \frac{\Delta p}{\Delta t} = \frac{\Delta(mv)}{\Delta t} = \frac{\Delta m(v-0)}{\Delta t} \xlongequal{m=\rho V} \frac{\rho \Delta V v}{\Delta t} \xlongequal{V=AL} \frac{\rho A \Delta L v}{\Delta t} \xlongequal{\Delta L = v\Delta t} \rho A v^2$$
$$\Rightarrow \quad F = 1.2 \times 10 \times 30^2 = 10800 \text{ N}$$
□

Example 6.4 Given the variation with time of the momentum of a body as shown in the p-t graph, check yourself that the variation of force acting and the variation of the object's acceleration should be plotted as shown in the F-t graph and the a-t graph.



Example 6.5 An object of mass 70 g is initially at rest. A force that varies with time is exerted on the object. The graph shows the how the force varies during the time of impact. What is the final velocity of the object?



✍ area under F-t graph gives change in momentum

$$\Delta p = \frac{1}{2} \times 50 \times 35 \times 10^{-3} = 0.875 \text{ kg m s}^{-1}$$

final velocity: $v = \dfrac{\Delta p}{m} = \dfrac{0.875}{70 \times 10^{-3}} = 12.5 \text{ m s}^{-1}$ □

6.1.3 principle of momentum conservation

change in an object's momentum is given by: $\Delta p = F\Delta t$ [37]

in particular, if there is zero net force, then object's momentum stays constant

this idea can be generalised to a system of objects

there are external forces from outside and internal forces between objects within the system

let's take two mutually interacting objects within the system, say A and B

by Newton's 3rd law, force on A by B is equal and opposite to force on B by A

change in A's momentum by B is then equal and opposite to change in B's momentum by A

change in total momentum of A and B due to each other is therefore cancelled out

therefore, for the system as a whole, effect of internal forces always cancel out

change of total momentum of the system would only depend on net external force [38]

---

[37] The relation $\Delta p = F\Delta t$ is valid if we are dealing with a *constant* force. If the object is acted by a varying force, then the change in its momentum is given by: $\Delta p = \int F \mathrm{d}t$.

[38] A more rigorous derivation goes as follows.

Let's consider a system of point objects m_i, each experiences a resultant force F_i where F_i can come from some external source or another object j within the system: $F_i = F_{i,\text{ext}} + \sum_j F_{i,j}$.

Summing over all objects, we can write: $\sum_i F_i = \sum_i F_{i,\text{ext}} + \sum_{i,j} F_{i,j}$

For each pair i and j, the action-reaction principle suggests that the mutual interaction between the two are equal but opposite: $F_{i,j} = -F_{j,i}$, so $\sum_{i,j} F_{i,j} = 0$. Therefore, $\sum_i F_i = \sum_i F_{i,\text{ext}}$.

Multiply both sides by Δt, we can write: $\sum_i F_i \Delta t = \sum_i F_{i,\text{ext}} \Delta t$. Note that $\sum_i F_i \Delta t = \sum_i \Delta p_i$ gives the change in total momentum of system, so this shows the change of total momentum is determined by the net external force: $\boxed{\left(\sum_i F_{i,\text{ext}}\right) \Delta t = \sum_i \Delta p_i}$

if there is no net external force, then no change in total momentum

> for any closed system where net external force is zero, the total momentum of the system remains constant, this is called the **principle of momentum conservation**

Example 6.6 A uranium-238 nucleus disintegrates, emitting an α-particle of mass 4u and producing a thorium-234 nucleus of mass 234u. The uranium nucleus is initially at rest. (a) What is the ratio of the velocities of the product particles $\dfrac{v_\alpha}{v_{\text{Th}}}$? (b) Explain why the α-particle and the thorium nucleus must be emitted in opposite directions.

✍ no external force involved during decay, so momentum is conserved

zero initial momentum means total momentum of α-particle and thorium nucleus is zero

α-particle and thorium nucleus must carry equal but opposite momenta

equal momentum $\quad\Rightarrow\quad m_\alpha v_\alpha = m_{\text{Th}} v_{\text{Th}} \quad\Rightarrow\quad \dfrac{v_\alpha}{v_{\text{Th}}} = \dfrac{m_{\text{Th}}}{m_\alpha} = \dfrac{234}{4} = 58.5$

opposite momentum so they move off in exactly opposite directions □

6.2 collision problems

for two bodies colliding together, external forces are negligible during the time of contact

hence total momentum is considered to be conserved for any collision process

6.2.1 collision in one dimension

suppose two masses m_1 and m_2 are restricted to move in one dimension only

they move at initial velocities u_1 and u_2 before they collide

after collision, their velocities become v_1 and v_2, as depicted in the diagram

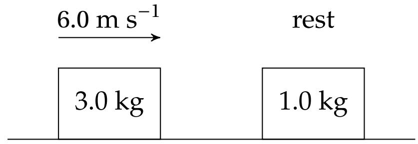
before collision after collision

total momentum is conserved, so: $\boxed{m_1 u_1 + m_2 u_2 = m_1 v_1 + m_2 v_2}$

➢ recall the vector nature of momentum and velocity

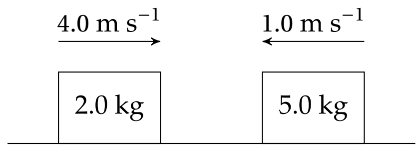
we normally choose objects moving to the right to have positive momentum/velocity

then object travelling to the left would have negative momentum/velocity



Example 6.7 A 3.0 kg mass moving at 6.0 m s^{-1} has a head-on collision with a stationary 1.0 kg mass. The two masses stick together on impact. What is the final velocity of the two masses?

✍ $$m_1u_1 + \cancelto{0}{m_2u_2} = (m_1+m_2)v \quad \Rightarrow \quad 3.0\times 6.0 = (3.0+1.0)\times v \quad \Rightarrow \quad v = 4.5\text{ m s}^{-1}$$ □



Example 6.8 A 2.0 kg mass moving at 4.0 m s^{-1} collides head on with a 5.0 kg mass moving at 1.0 m s^{-1}. After the collision, speed of the 5.0 kg mass is unchanged but its direction is reversed. What is the velocity of the 2.0 kg mass after the collision?

✍ $$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2 \Rightarrow 2.0\times 4.0 + 5.0\times(-1.0) = 2.0\times v_1 + 5.0\times 1.0 \Rightarrow v_1 = -1.0\text{ m s}^{-1}$$

minus sign means the 2.0 kg mass reverses direction after collision □

6.2.2 elastic & inelastic collisions

elastic collisions

collision can be either elastic or inelastic[39]

> for **elastic collisions**, there is no loss of kinetic energy

we hereby derive a condition that must be satisfied by two objects colliding elastically

since momentum and kinetic energy are both conserved, we can write two equations[40]

$$\boxed{m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2}$$

$$\boxed{\frac{1}{2}m_1u_1^2 + \frac{1}{2}m_2u_2^2 = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2}$$

rearranging both equations, we have

$$m_1(u_1 - v_1) = m_2(v_2 - u_2) \tag{1}$$

$$m_1(u_1^2 - v_1^2) = m_2(v_2^2 - u_2^2) \tag{2}$$

[39]Here we assume you already have some knowledge about *kinetic energy*. Kinetic energy of a moving body is given by the formula: $E_k = \frac{1}{2}mv^2$. You might have learned about it in an GCSE course or elsewhere. We will talk about kinetic energy in §7.2.

[40]For simplicity, we consider two-body collision in one dimension only, that is the two bodies move along the same straight line before and after the collision. For a two-body collision problem in two dimension, the conservation of momentum can be broken into two independent component equations.

equation (4) can be further rewritten as

$$m_1(u_1 - v_1)(u_1 + v_1) = m_2(v_2 - u_2)(v_2 + u_2) \tag{2'}$$

comparing equation (2′) to equation (1), one has: $u_1 + v_1 = v_2 + u_2$

rearrange the equation, we find: $\boxed{v_2 - v_1 = u_1 - u_2}$

both side of the equation now represent a *relative speed* between the two colliding bodies

> for an elastic collision process between two bodies, the relative velocity of separation after collision equals the relative velocity of approach before collision

inelastic collisions

> for **inelastic collisions**, part of kinetic energy is lost due to change in object's shape

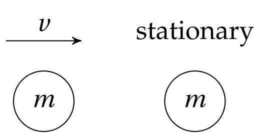
➢ for an inelastic process, the following will hold:
  - K.E. after collision is less than K.E. before collision
  - relative speed after collision is less than relative speed before collision

brief summary

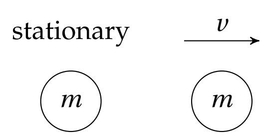
discussions on elastic and inelastic collisions are summarised in the table below

	elastic collision	inelastic collision
conservation of momentum	✓	✓
conservation of kinetic energy	✓	✗
conservation of total energy	✓	✓
relative speed stays unchanged	✓	✗

Example 6.9 A sphere of mass m moves on a smooth horizontal surface at speed v and collides *elastically* with an identical ball at rest. What are the final velocities of the two spheres?



before collision



after collision

$$\begin{cases} mv = mv_1 + mv_2 & \text{(momentum conservation)} \\ v = v_2 - v_1 & \text{(relative speed unchanged)} \end{cases} \Rightarrow \begin{cases} v_1 &= 0 \\ v_2 &= v \end{cases}$$

the two spheres simply exchange velocities during the collision, as shown in diagram[41] □

Example 6.10 A 4.2 kg mass A and a 1.5 kg mass B are travelling towards each other on a frictionless horizontal plane. Mass A and B move at 5.40 m s^{-1} and 1.80 m s^{-1} respectively before they strike, as shown below. Mass B moves to the right at 4.50 m s^{-1} after the collision, (a) find the velocity of A after the impact, and (b) suggest whether the collision is elastic.

before collision after collision

momentum conserved: $m_A u_A + m_B u_B = m_A v_A + m_B v_B$

$$4.2 \times 5.40 + 1.5 \times (-1.80) = 4.2 \times v + 1.5 \times 4.50 \quad \Rightarrow \quad v_A = 3.15 \text{ m s}^{-1}$$

K.E. before: $E_{k,i} = \frac{1}{2} m_A u_A^2 + \frac{1}{2} m_A u_B^2 = \frac{1}{2} \times 4.2 \times 5.40^2 + \frac{1}{2} \times 1.5 \times 1.80^2 \approx 63.7$ J

K.E. after: $E_{k,f} = \frac{1}{2} m_A v_A^2 + \frac{1}{2} m_A v_B^2 = \frac{1}{2} \times 4.2 \times 3.15^2 + \frac{1}{2} \times 1.5 \times 4.50^2 \approx 36.0$ J

there is K.E. loss, so collision is inelastic

alternatively, we can compare the relative speed before and after the collision

$$u_A - u_B = 5.40 - (-1.80) = 7.20 \text{ m s}^{-1} \qquad v_B - v_A = 4.50 - 3.15 = 1.35 \text{ m s}^{-1}$$

relative speed changed after the collision, so collision must be inelastic □

6.2.3 collision in two dimensions

when objects collide on a horizontal plane, they can possibly move off in any direction

negligible net external force is present, total momentum is still conserved

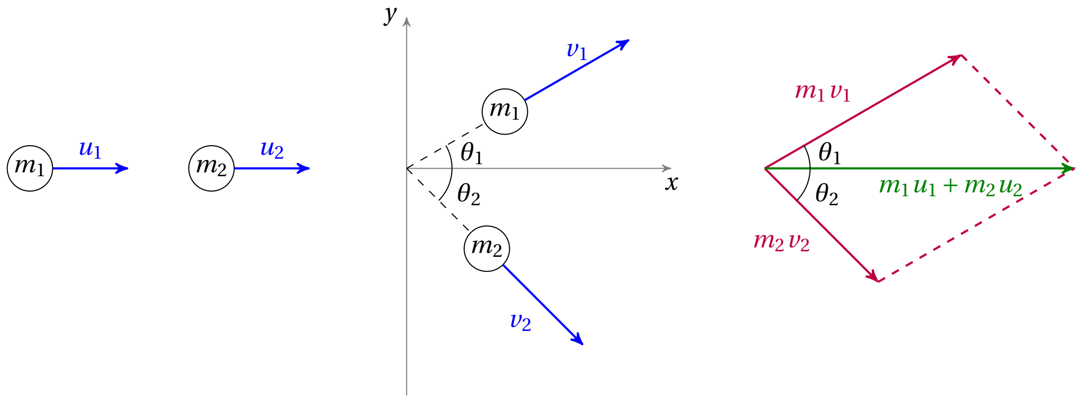
recall that momentum is a vector quantity, so momentum should be conserved in any direction

let's consider the collision between two masses m_1 and m_2

for simplicity, assume their initial velocities u_1 and u_2 are in same direction

final velocities v_1 and v_2 after the collision are shown

[41] If two objects of equal mass collide elastically with one another, one can actually show that their velocities would exchange regardless of their initial velocities.



before collision after collision momentum conservation

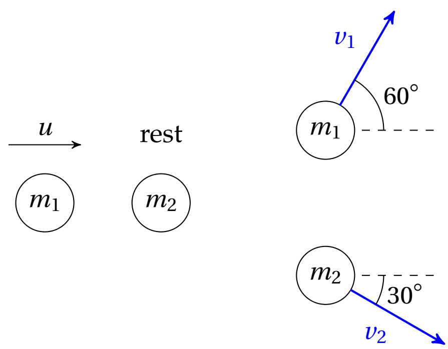
equations of momentum conservation can be written for two perpendicular directions

$$m_1 u_1 + m_2 u_2 = m_1 v_1 \cos\theta_1 + m_2 v_2 \cos\theta_2 \qquad \text{(in } x\text{-direction)}$$

$$0 = m_1 v_1 \sin\theta_1 - m_2 v_2 \sin\theta_2 \qquad \text{(in } y\text{-direction)}$$

one can also construct a vector triangle to transform the problem into a geometry problem

Example 6.11 A ball of mass $m_1 = 1.0$ kg travelling with a speed of $u = 6.0$ m s^{-1} in the x-direction strikes a stationary ball of mass $m_2 = 2.0$ kg. The direction of the balls' velocities v_1 and v_2 after the collision are shown in the diagram. Find v_1 and v_2.



before collision after collision

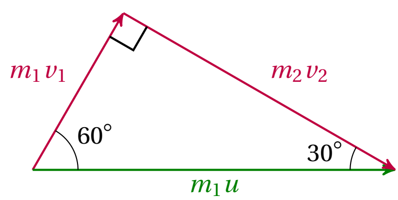
✍ start with momentum conservation equations:

$$\begin{cases} m_1 u = m_1 v_1 \cos 60^\circ + m_2 v_2 \cos 30^\circ \\ 0 = m_1 v_1 \sin 60^\circ - m_2 v_2 \sin 30^\circ \end{cases}$$

$$\begin{cases} 1.0 \times 6.0 = 1.0 \times v_1 \times \dfrac{1}{2} + 2.0 \times v_2 \times \dfrac{\sqrt{3}}{2} \\ 0 = 1.0 \times v_1 \times \dfrac{\sqrt{3}}{2} - 2.0 \times v_2 \times \dfrac{1}{2} \end{cases}$$

simplify and solve the equations:

$$\begin{cases} \dfrac{1}{2} v_1 + \sqrt{3} v_2 = 6 \\ v_2 = \dfrac{\sqrt{3}}{2} v_1 \end{cases} \Rightarrow \begin{cases} v_1 = 3.0 \text{ m s}^{-1} \\ v_2 \approx 2.6 \text{ m s}^{-1} \end{cases}$$



one can also draw and use the vector triangle

for this question, this happens to be a right-angled triangle, so things become much easier

$$\begin{cases} m_1 v_1 = m_1 u \cos 60^\circ \\ m_2 v_2 = m_1 u \cos 30^\circ \end{cases} \Rightarrow \begin{cases} 1.0 \times v_1 = 1.0 \times 6.0 \times \dfrac{1}{2} \\ 2.0 \times v_2 = 1.0 \times 6.0 \times \dfrac{\sqrt{3}}{2} \end{cases} \Rightarrow \begin{cases} v_1 = 3.0 \text{ m s}^{-1} \\ v_2 \approx 2.6 \text{ m s}^{-1} \end{cases} \qquad \square$$

Chapter 7
Work & Energy

In the previous chapter, we studied the accumulative effect of a force over a period of time and have seen how this gives rise to the idea of impulse and momentum. In this chapter, we consider the effect of a force over a certain displacement, and you will learn how this is related to the concept of work done and energy.

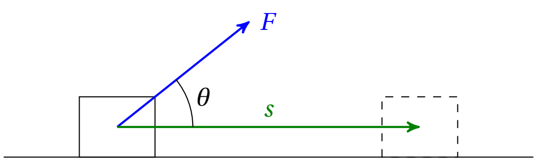
Energy is a concept central to all of physical sciences. The entire universe is made up by energy and matter. In this section we study the energy changes during various physical processes.

7.1 work

work done by a force is defined as the product of the force and the displacement moved out in the direction of the force: $\boxed{W = Fs}$

- unit for work done: $[W] = [F][s] = \mathrm{N} \cdot \mathrm{m} \equiv \mathrm{J}$ (joule)

  if a one newton force makes an object move out by one metre, then it does work of one **joule**
- if force acts at angle θ to the displacement travelled, then: $\boxed{W = Fs\cos\theta}$



- work is a *scalar* quantity[42], i.e., it has no direction
- work done can be either positive or negative

  resistive forces, such as friction and air drag, act in the opposite direction to motion

[42]Although work is defined as the product of two vectors, work carries no information about direction. This vector product is called as a *scalar product* or a *dot product*, which can be written explicitly as: $W = \vec{F} \cdot \vec{s} = |F||s|\cos\theta$. You might have seen this operation in the A-Level course in Mathematics.

so work done by resistive forces is negative[43], we say this is work *against* resistance

the minus sign will be crucial in energy calculations in later discussions

➢ a gas can do work to/against the surroundings

if pressure stays constant, then work by/on gas: $W = F\Delta s = pA\Delta s \quad \Rightarrow \quad \boxed{W_{\text{gas}} = p\Delta V}$

➢ work done by a varying force[44] is found by integration or using a F-s graph

if force varies with position, its change over small displacements is still considered small

work done over an infinitesimal displacement is therefore $\mathrm{d}W = F\mathrm{d}s$

integrate from initial position to final position, total work done is given by: $\boxed{W = \int_i^f F\mathrm{d}s}$

if one plots force against displacement, then **area under F-s curve gives work done**

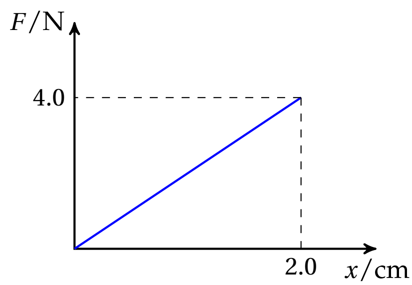
Example 7.1 A 20 N force is applied at 60° to the horizontal to move a 1.0 kg object at a constant speed of 2.0 m s^{-1} for 30 s. How much work is done by the force?

✍ $$W = Fs\cos\theta = Fvt\cos\theta = 20 \times 2.0 \times 30 \times \cos 60° \quad \Rightarrow \quad W = 600\text{ J}$$ □

Example 7.2 A piston in a gas pump has an area of 600 cm^2. During one stroke, the pump moves a distance of 30 cm against a constant pressure of 8000 Pa. How much work is done?

✍ $$W = p\Delta V = pA\Delta s = 8000 \times 600 \times 10^{-4} \times 30 \times 10^{-2} \quad \Rightarrow \quad W = 144\text{ J}$$ □

Example 7.3 When a spring is compressed by 2.0 cm, the force applied increases uniformly from zero to 4.0 N. How much work is done by this force?



✍ F-x graph for the force is plotted as shown

work to compress spring equals area under F-x graph:

$$W = \frac{1}{2} \times 4.0 \times 2.0 \times 10^{-2} = 0.040\text{ J}$$ □

7.2 types of energies

Energy is something acquired by an object that enables it to do work. A moving vehicle, water stored in a reservoir, a compressed spring, separated magnets, all of these objects can do work to other objects. In this section, we will look at various situations where work on an object causes a change in some form of energy.

[43]You might take $\theta = 180°$, then $\cos\theta = -1$, giving rise to a negative work done.

[44]The equation $W = Fs$ is valid only if the force acting is constant.

Similarly, the equation $W_{\text{gas}} = p\Delta V$ holds for constant pressure processes only.

7.2.1 kinetic energy

suppose a constant force F is acting over a distance s, we have:

$$W = Fs \overset{F=ma}{=\!=\!=} mas \overset{v^2=u^2+2as}{=\!=\!=\!=} m\frac{v^2-u^2}{2} \quad \Rightarrow \quad W = \frac{1}{2}mv^2 - \frac{1}{2}mu^2$$

this shows work is transformed into change in some quantity associated with object's motion

this is recognised as the gain in kinetic energy of the object: $\Delta E_k = \frac{1}{2}mv^2 - \frac{1}{2}mu^2$

> **kinetic energy** (K.E.) is the energy possessed by an object due to its motion

➢ an object of mass m moving with speed v has K.E.: $\boxed{E_k = \frac{1}{2}mv^2}$

Example 7.4 Estimate the kinetic energy of a running man.

✍ suppose the man has a mass of 75 kg and is running at 5 m s^{-1} (any reasonable value will do)

$$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times 75 \times 5^2 \approx 940 \text{ J}$$ □

7.2.2 gravitational potential energy

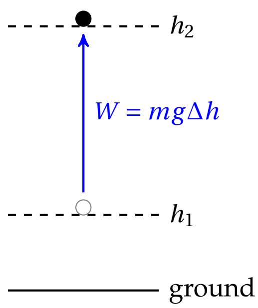
consider a body being slowly pulled from a height of h_1 to h_2

work done for this process is: $W = Fs = mg\Delta h = mgh_2 - mgh_1$

this shows work is transformed into change in some quantity associated with object's position

we say this is the gain in gravitational potential energy:

$$\Delta E_p = mgh_2 - mgh_1$$



> **gravitational potential energy** (G.P.E.) is the energy possessed by a body due to its position in a gravitational field

➢ a body of mass m at a height of h has G.P.E.: $\boxed{E_p = mgh}$

➢ G.P.E. is a *relative* quantity, only its change is important in physical processes [45]

[45] The formula $E_p = mgh$ implies that we have defined G.P.E. at the ground level to be zero. But this is purely conventional. Here I would like to point out that one can freely choose any zero potential energy level as he/she wishes, but no matter what point is picked as reference, we will always agree on the quantity of physical significance, that is, the *change* in G.P.E between two fixed points.

7.2.3 elastic potential energy

let's now consider a spring being stretched or compressed

work is done by external forces to cause the change in shape

this becomes of the elastic energy stored in the body

> **elastic potential energy**, also called **strain energy**, is the energy possessed by an elastic body due to deformation

this topic will be revisited in details in §8.1.2

7.2.4 other types of energies

apart from those we have mentioned above, there are many other types of energies

- *electric potential energy*: energy of a charged object due to its position in an electric field
- *chemical energy*: ability to do work due to potential energy between atoms and molecules
- *nuclear energy*: ability to do work due to potential energy of subatomic particles in the nuclei
- *internal energy*: sum of random kinetic and potential energies of molecules in a substance
- *electromagnetic energy*: energy carried by light/electromagnetic waves

7.2.5 work & energy transformations

from previous discussions, we have seen doing work is a way of transferring energy

for examination purposes[46], we can say the following:

> the change in the total energy of an object equals the net work done by all external forces (excluding those associated with potential energies): $\boxed{W = \Delta E}$

Example 7.5 A racing car of 800 kg starts off from rest. If the driving force is 5000 N, and the car experiences a constant resistive force of 1500 N, what is its speed after it has travelled 50 m?

✍ gain in K.E. equals work by driving force plus *negative* work against resistance

$$W_{\text{total}} = \Delta E_k \quad\Rightarrow\quad Fs - fs = \frac{1}{2}mv^2 - 0 \quad\Rightarrow\quad (5000-1500)\times 50 = \frac{1}{2}\times 800\times v^2 \quad\Rightarrow\quad v \approx 20.9 \text{ m s}^{-1} \quad \square$$

Example 7.6 A concrete cube of side 0.50 m and density 2400 kg m^{-3} is lifted 4.0 m by a crane. How much work is done?

[46] More rigorous discussions (which go beyond the syllabus) are given in §7.3.

✍ work by crane equals gain in G.P.E. of cube

$$W = \Delta E_p \quad \Rightarrow \quad W = mg\Delta h = \rho V g \Delta h = 2400 \times 0.50^3 \times 9.81 \times 4.0 \approx 1.18 \times 10^4 \text{ J}$$ □

7.3 conservation of energy

7.3.1 work-energy theorem (⋆)

more generally, net work done due to several forces $F_1, F_2, \ldots$, acting on an object is

$$W_{\text{total}} = \sum W_i = \sum \left(\int_i^f F_i \mathrm{d}s \right) = \int_i^f \left(\sum F_i \right) \mathrm{d}s = \int_i^f F_{\text{net}} \mathrm{d}s = \int_i^f ma \mathrm{d}s$$

recall in kinematics, $a = \dfrac{\mathrm{d}v}{\mathrm{d}t}$ and $\mathrm{d}s = v\mathrm{d}t$, so we have

$$W_{\text{total}} = \int_i^f m \frac{\mathrm{d}v}{\mathrm{d}t} v \mathrm{d}t = \int_i^f mv \mathrm{d}v$$

we are integrating over velocity from initial value u to final value v, so we find

$$W_{\text{total}} = \int_i^f mv \mathrm{d}v = \frac{1}{2} mv^2 \Big|_u^v = \frac{1}{2} mv^2 - \frac{1}{2} mu^2$$

identify kinetic energy $E_k = \dfrac{1}{2} mv^2$, now we come to the **work-energy theorem**

> the net work done on a body equals the change in the body's kinetic energy: $\boxed{W_{\text{total}} = \Delta E_k}$

7.3.2 conservative forces & potential energy (⋆)

if work by a force does not depend on which specific path is taken, this force is **conservative**

put differently, when an object is moved from one place to another under a conservative force, work done depends on initial and final positions only

we can split W_{total} into two parts: contributions from conservative and non-conservative forces

for conservative part, let's define: $\Delta E_p = -W_c = -\displaystyle\int_i^f F_c \mathrm{d}s$

this means work by conservative forces can be interpreted as change in its **potential energy**

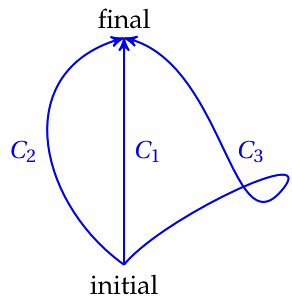
let's call the sum of an object's kinetic and potential energy its **mechanical energy**: $E_m = E_k + E_p$

work-energy theorem can then be rewritten: $\Delta E_k = W_{\text{total}} = W_c + W_{\text{nc}} = -\Delta E_p + W_{\text{nc}}$

rearrange we have: $W_{\text{nc}} = \Delta E_k + \Delta E_p \quad \Rightarrow \quad \boxed{W_{\text{nc}} = \Delta E_m}$

> the change in mechanical energy equals the net work done by all non-conservative forces

gravitational potential energy revisited ($\star$)



work by gravitational force is path independent

for example, same work is done via path C_1, C_2 and C_3

so gravitational force is conservative, G.P.E. can be defined

as a consequence, gain or loss in G.P.E. only depends on the difference in initial and final position of the object

if there is no other conservative force acting on a body apart from gravitational force, it follows that change in total energy of an object equals the net work done by all forces excluding gravity

7.3.3 conservation of energy

the law of conservation of energy states that energy cannot be created or destroyed, but can only transform from one form into another while the total amount is always constant

in absence of any non-conservative force, sum of a body's kinetic energy and potential energies, or the total mechanical energy, is constant, this is also a law of conservation

Example 7.7 For an object falling from rest due to gravity, if air resistance is negligible, what is its speed when it has fallen through a distance of h?

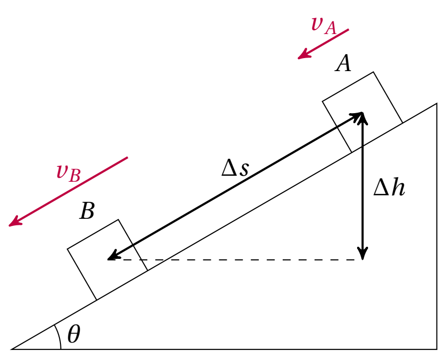
✍ $$\text{G.P.E loss} = \text{K.E gain} \quad\Rightarrow\quad mgh = \frac{1}{2}mv^2 \quad\Rightarrow\quad v = \sqrt{2gh}$$ □

Example 7.8 A marble is projected vertically upwards with an initial velocity u. The average resistive force acting is f. How do you determine the maximum height reached by the marble?

✍ K.E. loss = G.P.E. gain + energy loss due to resistance

$$\frac{1}{2}mu^2 - 0 = mgH_{\max} + fH_{\max} \quad\Rightarrow\quad H_{\max} = \frac{mu^2}{2(mg+f)}$$ □

Example 7.9 A box of mass m slides down along a slope that is inclined at an angle θ to the horizontal. There is a constant friction f acting on the box. When the box has moved through a distance of Δs down the slope from A to B, write down an equation relating its velocities v_A and v_B by applying the law of conservation of energy.



✍

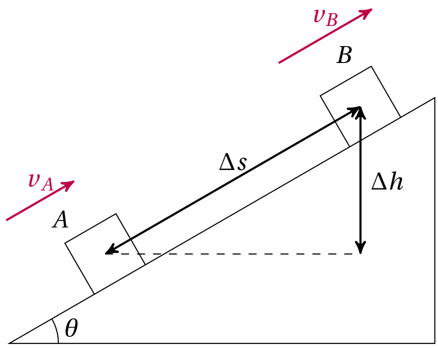
$$\begin{aligned}
\text{work done against friction} &= \text{change in total energy} \\
-W_f &= \Delta E_k + \Delta E_p \\
-fs &= \left(\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2\right) + (mgh_B - mgh_A) \\
-fs &= \left(\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2\right) - mg\Delta h \\
v_B^2 &= v_A^2 + 2g\Delta s\sin\theta - \frac{2fs}{m}
\end{aligned}$$

or equivalently, we can write

$$\begin{aligned}
\text{G.P.E loss} &= \text{K.E gain} + \text{energy loss due to friction} \\
mg\Delta h &= \left(\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2\right) + fs \\
v_B^2 &= v_A^2 + 2g\Delta s\sin\theta - \frac{2fs}{m}
\end{aligned}$$

note that the two alternative ways of thinking produce the same result □

Example 7.10 A slope is inclined at an angle θ to the horizontal. A box of mass m is pushed up the slope with a constant force F parallel to the slope, and the box experiences a constant frictional force f. When the box has moved through a distance of Δs along the slope from A to B, find an equation relating its velocities v_A and v_B.



✍

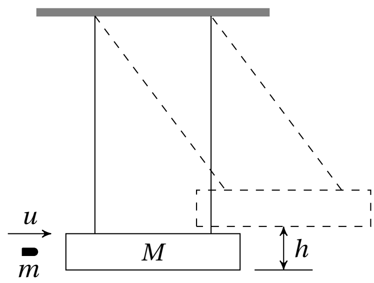
$$\begin{aligned}
\text{work by } F + \text{work against friction} &= \text{change in total energy} \\
W_F - W_f &= \Delta E_k + \Delta E_p \\
Fs - fs &= \left(\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2\right) + (mgh_B - mgh_A) \\
Fs - fs &= \left(\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2\right) + mg\Delta h \\
v_B^2 &= v_A^2 + \frac{2(F-f)s}{m} - 2g\Delta s\sin\theta
\end{aligned}$$

or equivalently, we can write

$$\begin{aligned}
\text{work by } F &= \text{K.E gain} + \text{G.P.E gain} + \text{energy loss due to friction} \\
Fs &= \left(\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2\right) + mg\Delta h + fs \\
v_B^2 &= v_A^2 + \frac{2(F-f)s}{m} - 2g\Delta s\sin\theta
\end{aligned}$$

again the two approaches produce the same expression for final velocity □

Example 7.11 A *ballistic pendulum* is a device used to measure the speeds of fast-moving bullets. It consists of a large block of wood of mass M, suspended from two long light strings. A bullet of mass m is fired into the block, and the block and bullet combination swings upward. If the centre of mass rises a vertical distance h, what is the initial speed u of the bullet?



✍ as bullet enters block, combined momentum is conserved:

$$mu = (M+m)v \quad \Rightarrow \quad v = \frac{mu}{M+m}$$

when the system swings upward, K.E. transforms into G.P.E. but total energy is conserved:

$$\frac{1}{2}(M+m)v^2 = (M+g)gh \quad \Rightarrow \quad v = \sqrt{2gh}$$

putting the two equations together, we find: $u = \left(1+\frac{M}{m}\right)\sqrt{2gh}$ □

7.4 power

to describe how fast work is done, we introduce the notion of power

power is defined as the work done per unit time: $\boxed{P = \frac{\Delta W}{\Delta t}}$

- unit of power: $[P] = \frac{[W]}{[t]} = \mathrm{J\,s^{-1}} = \mathrm{W}$ (watt)

  if one joule of work is done in one second, the power is one **watt**
- $P = \frac{\Delta W}{\Delta t}$ gives the *average* power during in a period of time Δt

  to find the *instantaneous* power at a particular moment, we have

$$P = \frac{\Delta W}{\Delta t} = \frac{F\Delta s}{\Delta t} \quad \Rightarrow \quad \boxed{P = Fv}$$

Example 7.12 There are 150 steps to the top of a tower, and the average height of each step is 25 cm. It takes a man of 72 kg two minutes to run up all the steps. What is his average power?

✍ $$P = \frac{\Delta E_p}{\Delta t} = \frac{mg\Delta h}{\Delta t} = \frac{72 \times 9.81 \times (150 \times 0.25)}{120} \quad \Rightarrow \quad P \approx 221\ \mathrm{W}$$ □

Example 7.13 A turbine is used to generate electrical power from the wind. Given that the blades of the turbine sweep an area of 500 m^2, the density of air is 1.3 kg m^{-3}, and the wind speed is 10 m s^{-1}. Assume no energy loss, find the power available from the wind.

✍ $$P = \frac{\Delta E_k}{\Delta t} = \frac{\frac{1}{2}\Delta m v^2}{\Delta t} = \frac{\frac{1}{2}\rho \Delta V v^2}{\Delta t} = \frac{\frac{1}{2}\rho A \Delta x v^2}{\Delta t} \quad \Rightarrow \quad P = \frac{1}{2}\rho A v^3$$

$$P = \frac{1}{2} \times 1.3 \times 500 \times 10^3 \approx 3.25 \times 10^5 \text{ W} \qquad \square$$

Example 7.14 A ship is cruising a a constant speed of 15 m s^{-1}. The total resistive force acting is 9000 N. What is the output power of this ship?

✍ constant speed so equilibrium between driving force and resistive force

$$P = Fv = fv = 9000 \times 15 \quad \Rightarrow \quad P = 1.35 \times 10^5 \text{ W} \qquad \square$$

Example 7.15 A car of mass 800kg accelerates from rest on a horizontal road. Suppose the engine provides a constant power of 24000 W, and the resistive force can be given by $f = 16v$, where v is the speed of the car in m s^{-1}. (a) What is the acceleration of the car when it is travelling at 15 m s^{-1}? (b) What happens to the car if it maintains this driving power?

✍ equation of motion for the car is: $F_{\text{net}} = F - f = ma \quad \Rightarrow \quad \frac{P}{v} - \alpha v = ma$

at $v = 15$ m s^{-1}: $\frac{24000}{15} - 16 \times 15 = 800a \quad \Rightarrow \quad a = 1.7$ m s^{-2}

as car's velocity v increases, driving force $F = \frac{P}{v}$ decreases, resistive force $f = \alpha v$ increases

so resultant force will decrease, the car will accelerate at a decreasing rate

eventually it reaches an equilibrium state where $F = f$, the car then travels at constant speed v_t

$$F = f \quad \Rightarrow \quad \frac{P}{v_t} = \alpha v_t \quad \Rightarrow \quad \frac{24000}{v_t} = 16 v_t \quad \Rightarrow \quad v_t \approx 38.7 \text{ m s}^{-1} \qquad \square$$

7.5 efficiency

efficiency of a system is given by: efficiency $= \frac{\text{useful energy output}}{\text{total energy input}}$, or $\boxed{\eta = \frac{W_{\text{useful}}}{W_{\text{total}}}}$

since $\Delta W = P\Delta t$, efficiency can also be evaluated in terms of power: $\boxed{\eta = \frac{P_{\text{useful}}}{P_{\text{total}}}}$

Example 7.16 A water pumping system uses 3.0 kW of electrical power to raise water from a well. The pump lifts 1500 kg of water per minute through a vertical height of 8.0 m. What is the efficiency of the sysem?

✍ $$\eta = \frac{\Delta E_p}{\Delta E_{\text{in}}} = \frac{\Delta mgh}{P_{\text{in}} \Delta t} = \frac{1500 \times 9.81 \times 8.0}{3000 \times 60} \quad \Rightarrow \quad \eta \approx 65.4\% \qquad \square$$

Example 7.17 Water flows into a turbine from a reservoir at a vertical distance of 70 m above. The water flows through the turbine at a rate of 2500 kg per minute. What is the output power of the turbine if it is 85% efficient?

✍ $$P_{\text{out}} = \eta P_{\text{in}} = \eta \frac{\Delta E_p}{\Delta t} = \eta \frac{\Delta mgh}{\Delta t} = 85\% \times \frac{2500 \times 9.81 \times 70}{60} \quad \Rightarrow \quad P_{\text{out}} \approx 2.43 \times 10^4 \text{ W}$$

Chapter 8

Solids

in this chapter, we study how a force changes the shape of an object

an important notion is the elasticity of materials

for **elastic** materials, when external force is removed, it can return to its original shape

if the material cannot restore to original shape, it is said to be **inelastic**, or **plastic**

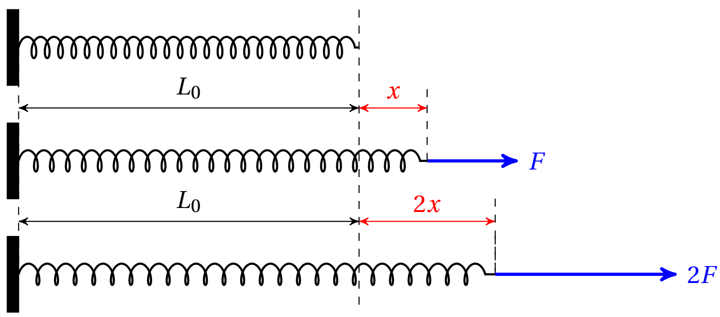
8.1 springs

8.1.1 Hooke's law

when a force F is applied to a spring, it is stretched from original length L_0 to some length L

extension of a spring, $x = L - L_0$, is dependent on the force applied

extension of an ideal spring is directly proportional to the load applied (within a certain range), this is called **Hooke's law**



- Hooke's law can be summarised by the equation: $\boxed{F = kx}$
- the proportionality constant k is called the **spring constant**, or **force constant**

  larger k means a greater force is required to extend the spring by same amount

  a spring with a large k is said to be *stiff*
- Hooke's law also holds if spring is being compressed

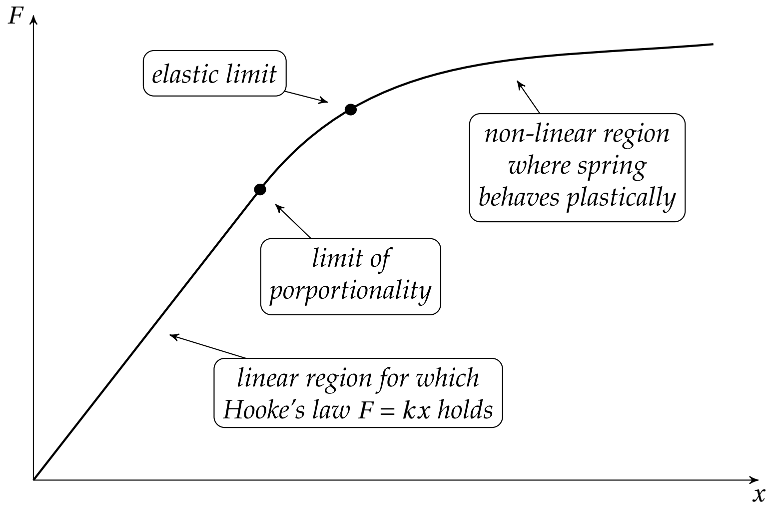
i.e., if a spring is pushed to have compression of x, we still have force applied $F = kx$

➢ linear relationship between F and x is only true up to a certain range

the limit at which Hooke's law no longer holds is called the **limit of proportionality**

➢ if load is too large, spring might be overstretched and no longer exhibits elastic behaviour

the point beyond which spring cannot return to original length is called the **elastic limit** [47]



force-extension graph for a typical spring under load

Example 8.1 A spring has a natural length of 20.0 cm. When a mass of 250 g is suspended from the spring, the new length of the spring is 26.0 cm. Find the spring constant.

✍ $$k = \frac{F}{x} = \frac{mg}{L - L_0} = \frac{0.250 \times 9.81}{(26.0 - 20.0) \times 10^{-2}} \quad \Rightarrow \quad k \approx 40.9\text{ N m}^{-1}$$ □

Example 8.2 A spring has a spring constant of 270 N m^{-1}. A mass of 1.2 kg is hung from the spring. When the mass is released from a position where the spring has an extension of 5.0 cm, what is the acceleration of the mass?

✍ $$F_{\text{net}} = F - mg = kx - mg = ma \quad \Rightarrow \quad a = \frac{kx - mg}{m} = \frac{270 \times 5.0 \times 10^{-2} - 1.2 \times 9.81}{1.2} \approx 1.44\text{ m s}^{-2}$$ □

8.1.2 elastic potential energy in a spring

to stretch or compress a spring, work must be done

this becomes of *elastic potential energy* stored in the spring

[47]The elastic limit of a spring and the limit of proportionality are two different but always confused concepts. These two limits are usually very close to one another, but they are conceptually different.

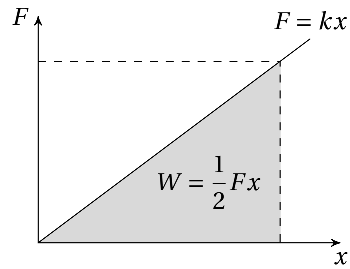
note that force in spring varies as spring is stretched

to find work in stretching a spring by x, we compute area under F-x graph[48], which is a right-angled triangle

$$W = \frac{1}{2}Fx = \frac{1}{2}kx^2$$

so elastic potential energy stored in a spring is:

$$\boxed{E_p = \frac{1}{2}kx^2}$$



Example 8.3 A steel spring has a spring constant of 20 N cm^{-1}. How much work is needed to stretch it from an extension of 3.0 cm to an extension of 5.0 cm?

✍ work done needed equals the increase in elastic potential energy:

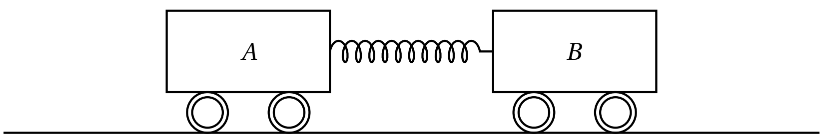
$$W = \Delta E_p = \frac{1}{2}kx_2^2 - \frac{1}{2}kx_1^2 = \frac{1}{2} \times 2000 \times (0.050^2 - 0.030^2) = 1.6 \text{ J}$$ □

Example 8.4 A trolley of 400 g can travel freely along a horizontal surface. It is pushed against a spring buffer. Suppose the spring is initially compressed by 5.0 cm under a 20 N force. When the trolley is released, it accelerates until it becomes detached. What is the trolley's final speed?

✍ elastic potential energy in spring transforms into kinetic energy of trolley

$$\frac{1}{2}Fx = \frac{1}{2}mv^2 \quad \Rightarrow \quad \frac{1}{2} \times 20 \times 0.050 = \frac{1}{2} \times 0.40 \times v^2 \quad \Rightarrow \quad v \approx 1.58 \text{ m s}^{-1}$$ □

Example 8.5 The same spring is now set between two trolleys A and B of mass 400 g and 600 g. Initially the spring is again compressed by 5.0 cm under a force of 20 N. After both trolleys are released, what are their final speeds?



✍ elastic potential energy in spring transforms into kinetic energy of the two trolleys:

$$\frac{1}{2}Fx = \frac{1}{2}m_A v_A^2 + \frac{1}{2}m_B v_B^2 \quad \Rightarrow \quad 20 \times 0.050 = 0.40 v_A^2 + 0.60 v_B^2$$

total momentum for trolley A and B as a whole is conserved:

$$m_B v_B - m_A v_A = 0 \quad \Rightarrow \quad 0.40 v_A = 0.60 v_B$$

solving the simultaneous equations, we find: $v_A \approx 1.22 \text{ m s}^{-1}$, $v_B \approx 0.82 \text{ m s}^{-1}$ □

[48]Mathematically, we can also integrate over the total extension to find this work done $W = \int_0^x F\mathrm{d}x = \int_0^x kx\mathrm{d}x = \frac{1}{2}kx^2$, which of course gives the same result.

8.1.3 spring combinations

so far we have discussed the properties of a single spring

next we investigate how a set of springs respond to a given load

parallel springs

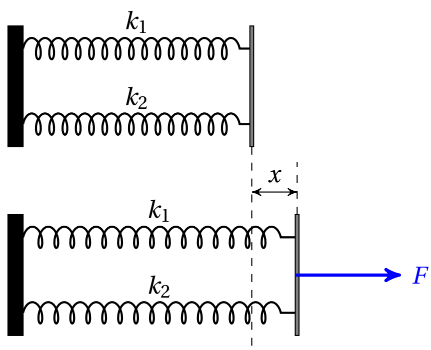
let's take two springs connected in parallel

when the combination is stretched under a load of F, extension in each spring should be the same:

$$x_1 = x_2 = x$$

the forces in each spring will in general be different, but sum of these must be equal to load:

$$F = F_1 + F_2$$



divide both sides by x, we have: $\dfrac{F}{x} = \dfrac{F_1}{x_1} + \dfrac{F_2}{x_2}$.

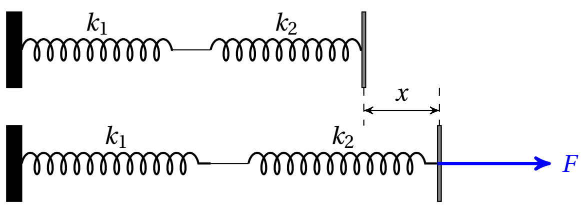
recall the Hooke's law, this becomes: $k = k_1 + k_2$

generalize for n springs in parallel connection, the combined spring constant is

$$\boxed{k = k_1 + k_2 + \cdots + k_n}$$

series springs

let's now take two springs in series



force in each spring is the same: $F_1 = F_2 = F$

but total extension is the sum of individual extensions: $x = x_1 + x_2$.

divide by the same F, we have: $\dfrac{x}{F} = \dfrac{x_1}{F_1} + \dfrac{x_2}{F_2}$.

for combined spring constant, we find: $\dfrac{1}{k} = \dfrac{1}{k_1} + \dfrac{1}{k_2}$

for n springs connected in series, the combined spring constant is therefore given by

$$\boxed{\frac{1}{k} = \frac{1}{k_1} + \frac{1}{k_2} + \cdots \frac{1}{k_n}}$$

some brief remarks

- if we set up n springs in parallel, we actually make it thicker

  it then requires a stronger force to stretch it by the same extension

  so we indeed see the combined spring constant is greater than that of any individuals
- if we set up n springs in series, we make it longer

  instead of pulling one spring at a time, the same force now stretches n springs simultaneously

  this gives rise to a greater total extension

  so the combined spring constant must be less than that of any individual

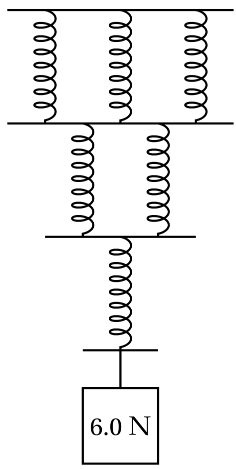
Example 8.6 (a) A spring with $k_1 = 20\ \text{N cm}^{-1}$ is connected in series with a second spring with $k_2 = 30\ \text{N cm}^{-1}$. When a force of 60 N is applied, what is the total extension of the combination? (b) The same two springs are now connected in parallel. When a force of 50 N is applied on the combination, what is the extension?

✍ for series connection: $k = \left(\frac{1}{k_1} + \frac{1}{k_1}\right)^{-1} = \left(\frac{1}{20} + \frac{1}{30}\right)^{-1} = 12\ \text{N cm}^{-1} \quad \Rightarrow \quad x = \frac{F}{k} = \frac{60}{12} = 5.0\ \text{cm}$

or sum up extension of each spring: $x = x_1 + x_2 = \frac{F}{k_1} + \frac{F}{k_2} = \frac{60}{20} + \frac{60}{30} = 5.0\ \text{cm}$

for parallel connection: $k = k_1 + k_2 = 20 + 30 = 50\ \text{N cm}^{-1} \quad \Rightarrow \quad x = \frac{F}{k} = \frac{50}{50} = 1.0\ \text{cm}$ □

Example 8.7 A set of identical springs are set up as shown. Each individual spring extends by 1.0 cm under a load of 1.0 N. Assume the limit of proportionality is not exceeded, what is the total extension for this combination when a load of 6.0 N is applied?



✍ for a single spring: $k = 1.0\ \text{N cm}^{-1}$

combined spring constant: $k_{\text{total}} = \left(\frac{1}{3k} + \frac{1}{2k} + \frac{1}{k}\right)^{-1}$

$$k_{\text{total}} = \left(\frac{1}{3.0} + \frac{1}{2.0} + \frac{1}{1.0}\right)^{-1} = \frac{11}{6} \approx 1.83\ \text{N cm}^{-1}$$

total extension: $x = \frac{F}{k_{\text{total}}} = \frac{6.0}{1.83} = 11\ \text{cm}$

alternatively, we can find and add extensions of each layer

in particular, the top layer withstands a force of 6.0 N shared by three springs, so each spring has a force of 2.0 N, extension is 2.0 cm

similarly, the other two layers extend by 3.0 cm and 6.0 cm respectively

hence, total extension: $x = 2.0 + 3.0 + 6.0 = 11\ \text{cm}$ □

8.2 stress, strain & Young modulus

from daily experience, it is easier to stretch a longer wire than a shorter one

same tensile force also produce greater effects on a thinner material than on a thicker one

so we need better quantities to describe the amount of deformation and the amount of action

to study a material's response to a tensile force, several new quantities are to be introduced

8.2.1 stress & strain

tensile strain is defined as the ratio of the extension of a wire to its natural length: $\boxed{\epsilon = \frac{x}{L}}$

tensile stress is defined as the force applied per unit cross-sectional area: $\boxed{\sigma = \frac{F}{A}}$

- strain is the ratio of two lengths, so it is unit free

  strain is usually expressed in terms of a percentage number
- units of stress: $[\sigma] = \text{N m}^{-2} = \text{Pa}$ (pascal)
- elastic behaviour of a wire is more or less like that of a spring

  Hooke's law says $F \propto x$, it then follows that $\frac{F}{A} \propto \frac{x}{L}$, so $\sigma \propto \epsilon$

  i.e., stress and strain should also be proportional to each other within certain limit

Example 8.8 A copper wire has a cross-sectional area of 1.5×10^{-6} m^2. The breaking stress of the wire is 2.0×10^8 Pa. Find the breaking force.

✍ $$\sigma = \frac{F}{A} \quad \Rightarrow \quad F = \sigma A = 2.0 \times 10^8 \times 1.5 \times 10^{-6} = 3.0 \times 10^2 \text{ N}$$ □

8.2.2 Young modulus

ratio of stress to strain of a material is called the **Young modulus**: $\boxed{E = \frac{\sigma}{\epsilon}}$

- unit of measurement: $[E] = \text{Pa}$ (pascal)
- Young modulus is a property of the material

  for the same material, Young modulus is a constant, no matter in what shape it takes i.e., it does not depend on the length or the cross section of the object

➢ typical value of Young's modulus for metals: $E_{\text{metal}} \sim 10^{11}$ Pa

➢ Young modulus is a measure of the stiffness of a material

to produce same strain, greater stress is required for a material with greater Young modulus

➢ Young modulus E is related to the force constant k

from $E = \dfrac{\sigma}{\epsilon} = \dfrac{F/A}{x/L} = \dfrac{FL}{xA}$, we rearrange to get: $F = \dfrac{EA}{L}x$

compare with Hooke's law, we can identify the force constant to be given by: $k = \dfrac{EA}{L}$

– $E \nearrow \Rightarrow k \nearrow$, stiffer material makes stiffer springs

– $A \nearrow \Rightarrow k \nearrow$, more difficult to stretch a thick spring (think about parallel springs)

– $L \nearrow \Rightarrow k \searrow$, easier to stretch a long spring (think about series springs)

Example 8.9 A 200 N tensile force is applied on a steel wire of 1.5 m, the wire extends by 5.0 mm. The diameter of the cross section is 0.60 mm. What is the Young modulus of the steel wire?

✍ $$E = \frac{\sigma}{\epsilon} = \frac{F/A}{x/L} = \frac{FL}{Ax} = \frac{200 \times 1.5}{\pi \times (0.30 \times 10^{-3})^2 \times 5.0 \times 10^{-3}} \quad \Rightarrow \quad E \approx 2.1 \times 10^{11} \text{ Pa} \qquad \square$$

Example 8.10 A copper wire of length 2.0 m is under a stress of 7.8×10^7 Pa. Given that the Young modulus of copper is 1.2×10^{11} Pa, what is (a) the strain of the wire, (b) the extension of the wire?

✍ strain: $\epsilon = \dfrac{\sigma}{E} = \dfrac{7.8 \times 10^7}{1.2 \times 10^{11}} \quad \Rightarrow \quad \epsilon = 6.5 \times 10^{-4} = 0.065\%$

extension: $x = \epsilon L = 6.5 \times 10^{-4} \times 2.0 \quad \Rightarrow \quad x = 1.3 \times 10^{-3}$ m $\square$

Example 8.11 Several blocks of steel are used to support a bridge. Each block has a height of 30 cm and a cross section of 15 cm × 15 cm. The steel block is designed to compress 2.0 mm when the maximum load is applied. Given that the Young modulus of steel is 2.1×10^{11} Pa, what is the maximum load that can be supported by one block?

✍ $$E = \frac{FL}{Ax} \quad \Rightarrow \quad F = \frac{EAx}{L} = \frac{2.1 \times 10^{11} \times (0.15 \times 0.15) \times 2.0 \times 10^{-3}}{0.30} \quad \Rightarrow \quad F \approx 3.15 \times 10^7 \text{ N} \qquad \square$$

Example 8.12 Two metal wires A and B are of the same length and they extend by the same amount under the same load. Given that Young modulus of wire A is twice of B, what is the ratio of their diameters?

✍ $$E = \frac{FL}{Ax} \quad \Rightarrow \quad A = \frac{1}{4}\pi d^2 = \frac{FL}{Ex} \quad \Rightarrow \quad d^2 \propto \frac{1}{E} \quad \Rightarrow \quad \frac{d_A}{d_B} = \sqrt{\frac{E_B}{E_A}} = \frac{1}{\sqrt{2}} \qquad \square$$

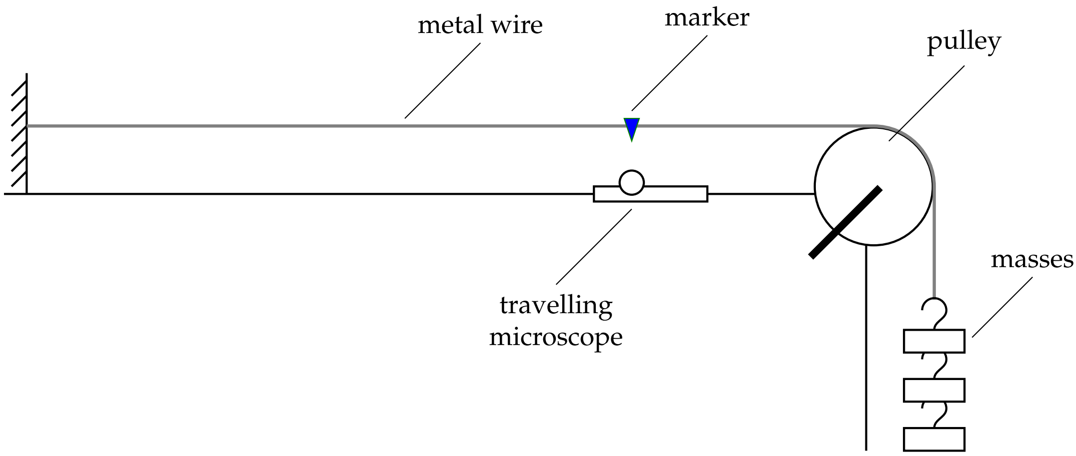
Example 8.13 A full-size crane is ten times greater than a model crane in all linear dimensions. If they are made of the same materials, what is the ratio of the cable's extension?

✍ same material means same density ρ and same Young's modulus E, so:

$$E = \frac{FL}{Ax} \quad \Rightarrow \quad x = \frac{FL}{AE} = \frac{mgL}{\pi r^2 E} = \frac{\rho V g L}{\pi r^2 E} \quad \Rightarrow \quad x \propto \frac{VL}{r^2} \propto \frac{l^3 l}{l^2} \propto l^2 \quad \Rightarrow \quad \frac{x_{\text{full}}}{x_{\text{model}}} = 10^2 = 100 \qquad \square$$

8.2.3 measurement of Young modulus

to measure Young's modulus of a metal wire, experimental setup can be laid out as shown



➢ method of data collection

– original length L (up to the marker) of the wire is measured with *metre rule*

– diameter d of the wire is measured with *micrometer*, then cross-sectional area is: $A = \frac{1}{4}\pi d^2$

– record mass m attached to the wire, then force applied is $F = mg$

– extension x of the wire is taken to be distance moved out by the marker

this can be measured with a *travelling microscope*[49] or a *vernier calliper*

➢ analysis of data

stress can be calculated by $\sigma = \frac{F}{A}$, and strain can be calculated by $\epsilon = \frac{x}{L}$

a graph of stress against strain can be plotted, a best fit curve can be drawn

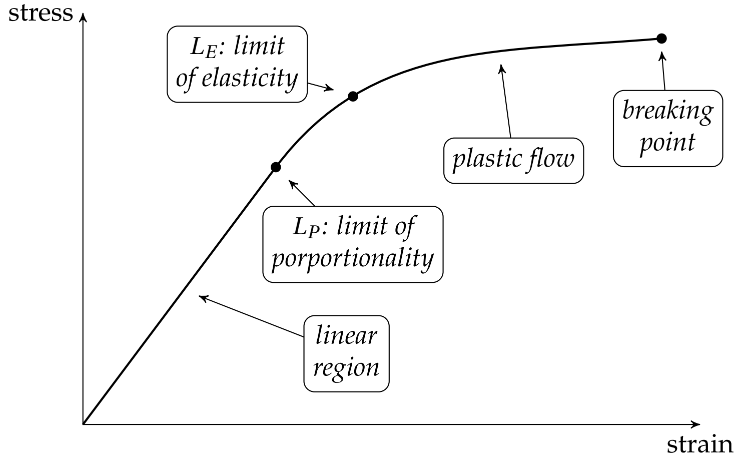
gradient of the straight-line section gives Young modulus

8.2.4 stress-strain curves

stress-strain curve for a material can be obtained using the methods in §8.2.3

in AS-Level, you are only supposed to know the behaviour of a metal under stress

[49] A travelling microscope is basically a microscope that can move back and forth along a rail. The position of the microscope can be varied by turning a screw. This position can be read off a vernier scale. So in short, a travelling microscope can be used to measure the change in length with a very high resolution (typically to a precision of 0.01mm or 0.02 mm).



stress-strain graph for a typical metal

- up to limit of proportionality L_E, stress is proportional to strain

  Young modulus can be given by the gradient of the curve before L_P
- consider the product of stress and strain: $\sigma \cdot \epsilon \sim \frac{F}{A}\frac{x}{L} \sim \frac{Fx}{AL} \sim \frac{W}{V}$

  so area under stress-strain curve gives the work done per unit volume to stretch the wire
- up to limit of elasticity L_E, metal wire can return to original length when it relaxes
- if a wire is stretched beyond the elastic limit, it follows a different path when force is removed

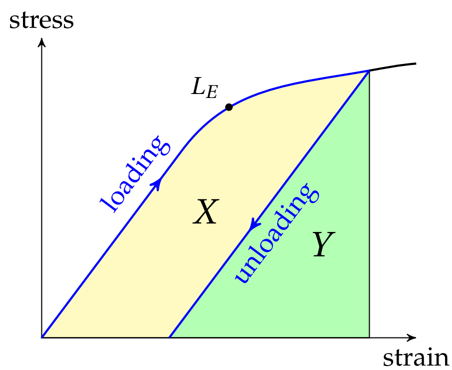
  recall area under stress-strain graph relates to energy

  work to stretch wire is given by $X + Y$

  energy goes out when wire contracts is Y

  the difference X gives energy loss during one cycle

  this energy difference becomes heat produced in wire



Chapter 9
Fluids

a fluid, such as a liquid or a gas, is a substance that has no fixed shape

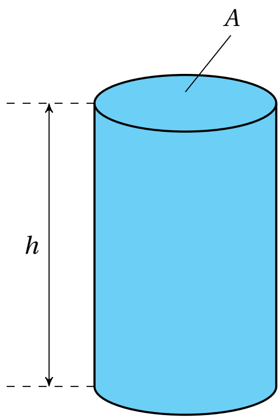
unlike a solid, a fluid can flow and yield easily under external force

in this chapter, we will study several aspects of a fluid

9.1 pressure in a fluid

at a depth of h below surface of a fluid, self-weight of the fluid could produce a pressure

$$p = \frac{F}{A} = \frac{mg}{A} = \frac{\rho V g}{A} \quad \Rightarrow \quad \boxed{p = \rho g h}$$



➢ pressure in a liquid depends on depth

for different positions in a liquid, as long as they are at same depth, pressure is the same

i.e., pressure does not depend on volume or shape of container

➢ atmospheric pressure also accounts for total pressure in a liquid

atmosphere presses on surface of a liquid, so total pressure at depth h is: $P = \rho g h + P_{atm}$

nevertheless, change in pressure still satisfies: $\Delta p = \rho g \Delta h$

Example 9.1 The atmospheric pressure is about 1.0×10^5 Pa. Given that the density of sea water is 1020 kg m^{-3}, what is the total pressure 50 m below the surface of the sea?

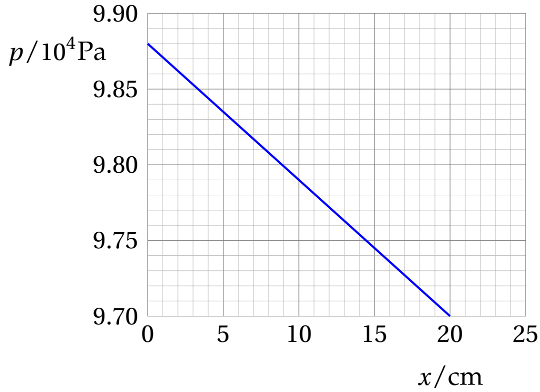
✍ $$P = P_{atm} + \rho g h = 1.0 \times 10^5 + 1020 \times 9.81 \times 50 \quad \Rightarrow \quad P \approx 6.0 \times 10^5 \text{ Pa}$$ □

Example 9.2 A vertical column of liquid of height 10 m contains both oil and water. The pressure due to the liquids at the bottom of the column is 89 kPa. Given that the density of water is 1000 kg m^{-3} and the density of the oil is 840 kg m^{-3}. What is the depth of the oil?

✍ $$P = P_{oil} + P_{water} = \rho_o g h_o + \rho_w g h_w$$

$$840 \times 9.81 \times x + 1000 \times 9.81 \times (10 - x) = 89 \times 10^3 \quad \Rightarrow \quad x = 5.8 \text{ m}$$ □

Example 9.3 The pressure p of a liquid in a container varies with the height x above the base of the container as shown. The total depth of the liquid is 20 cm. (a) What is the atmospheric pressure? (b) What is the density of the liquid?



✍ surface of liquid at height $x = 20$ cm, so

$$P_{\text{atm}} = 9.70 \times 10^4 \text{ Pa}$$

density of liquid: $\rho = \dfrac{\Delta p}{g\Delta h} = \dfrac{(9.88 - 9.70) \times 10^4}{9.81 \times 0.20} \quad \Rightarrow \quad \rho \approx 917 \text{ kg m}^{-3}$ □

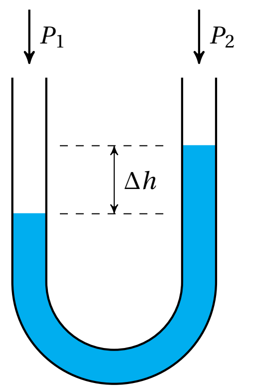
9.2 pressure meters

there are many types of instruments for pressure measurement

we would only focus on two: simple manometers and barometers [50]

they both use the fact that $\Delta p = \rho g \Delta h$ within a liquid

9.2.1 manometers



a **manometer** consists of a U-shaped tube filled with some liquid

any pressure difference between the two ends of the tube could cause a height difference between liquid levels

for the situation shown, at equilibrium, one has: $P_1 - P_2 = \rho g \Delta h$

if P_2 is a reference pressure, then P_1 can be calculated

➢ though any fluid can be used in a manometer, *mercury* is preferred because of its high density ($\rho_{\text{Hg}} = 1.36 \times 10^4 \text{ kg m}^{-3}$)

Example 9.4 A manometer is used to measure the pressure of a gas supply. Side A of the tube is connected to the gas pipe, and the other side B of the tube is open to the atmosphere. If the

[50]Some examples of many other types of pressure gauges include

- mechanical gauges based on metallic pressure-sensing elements
- electronic gauges based on piezo-resistive effect
- hot-filament ionization gauges based on ion currents from a gas

Those who are interested are welcome to research into their functions and principles.

mercury on side A is higher than on side B by 14 cm, what is the pressure of the gas? (density of mercury: 1.36×10^4 kg m^{-3}; atmospheric pressure: 1.01×10^5 Pa)

✍ $P_{atm} - P_{gas} = \rho g h \quad \Rightarrow \quad P_{gas} = 1.01 \times 10^5 - 1.36 \times 10^4 \times 9.81 \times 0.14 \quad \Rightarrow \quad P_{gas} \approx 8.23 \times 10^4$ Pa □

9.2.2 barometers

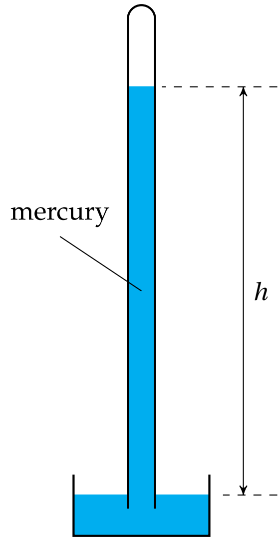
take a long glass tube and fill it with mercury

let it stand upside down in a basin

there is atmospheric pressure pushing down on surface of mercury, so a height of mercury is supported up the tube

we can then compute atmospheric pressure by: $P_{atm} = \rho g h$

this instrument makes a **mercury barometer**



Example 9.5 If a mercury barometer supports a height of 760 mm of mercury above the fluid level in the container, what is the atmospheric pressure? If water is used as the barometric liquid, what is the minimum length of the tube required for the same atmospheric pressure? (density of mercury: 1.36×10^4 kg m^{-3}; density of water: 1.00×10^3 kg m^{-3};)

✍ atmospheric pressure: $P_{atm} = \rho g h = 1.36 \times 10^4 \times 9.81 \times 0.760 \approx 1.01 \times 10^5$ Pa

if mercury is replaced by water: $h' = \dfrac{P_{atm}}{\rho' g} = \dfrac{1.01 \times 10^5}{1000 \times 9.81} \approx 10.3$ m □

9.3 upthrust

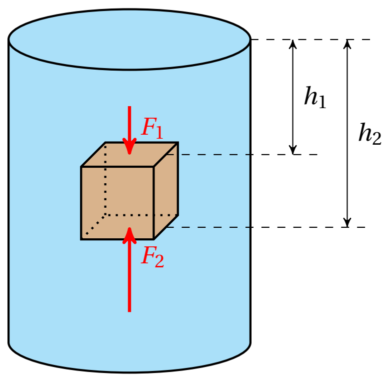
now consider a rectangular block immersed in a fluid

top and bottom surface are at different depths, so they experience different pressures

this gives rise to an overall upward force on the cylinder

this force is called the **upthrust**:

$$F_U = F_2 - F_1 = \rho g (h_2 - h_1) \times A \quad \Rightarrow \quad \boxed{F_U = \rho g V}$$



therefore upthrust exerted on an immersed object equals the weight of the fluid displaced

this is known as the **Archimedes' principle**

➢ origin of upthrust: pressure difference between top and bottom surfaces

➢ for an object of density ρ_o immersed in a liquid of density ρ_l

take force in downward direction to be positive, then resultant force acting is:

$$F_{net} = W - F_U = \rho_o g V - \rho_l g V = (\rho_o - \rho_l) g V$$

– if $\rho_o > \rho_l$, then $F_{net} > 0$, resultant force acts downwards, object will sink

– if $\rho_o < \rho_l$, then $F_{net} < 0$, resultant force acts upwards, object will rise

– if $\rho_o = \rho_l$, then $F_{net} = 0$, object is in equilibrium, it can float at that level

Example 9.6 A block of mass 80 g and volume 50 cm^3 is suspended from a string into water. When the block is fully immersed and kept at rest, what is the tension in the string?

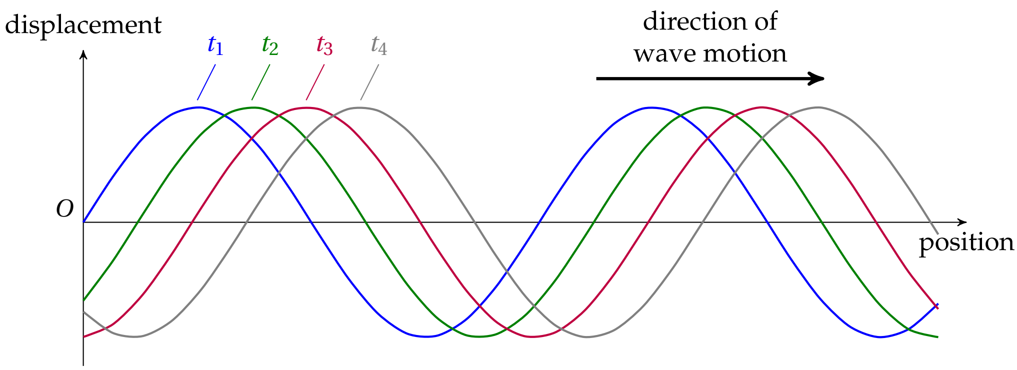
✍ $T + F_U = W \quad \Rightarrow \quad T = mg - \rho g V = 0.080 \times 9.81 - 1000 \times 9.81 \times 50 \times 10^{-6} \quad \Rightarrow \quad T \approx 0.29\ \mathrm{N}$ □

Chapter 10
Waves

Wave is a way of transferring and storing *energy* without the transport of matter. Most familiar examples are surface waves on water, sound waves, light waves. In the next two chapters, we are going to look at the basics of wave motion and some of the most important wave phenomena.

10.1 wave terminologies

wave motion is the propagation of disturbance from one place to another



wave pattern at different times ($t_1 < t_2 < t_3 < t_4$) as wave travels in space

to describe the wave motion and particle vibrations, we can define the following quantities:

- as a wave moves out, each point oscillates back and forth about their rest positions

  distance from a particle's equilibrium position is called **displacement** of the particle
- greatest displacement for a particle is called the **amplitude** (A)
- wave pattern repeats itself over a certain distance

  distance between two adjacent points undergoing exactly same motion is the **wavelength** (λ)

  one can think of wavelength as crest-to-crest distance, trough-to-trough distance, etc.[51]

[51]For now, we take for granted that a wave is transverse. There are also longitudinal waves for which terms like crest and trough do not apply. We will get into that in §10.2.2.

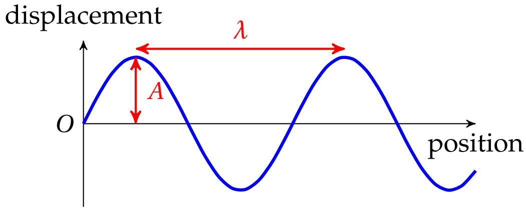
- each point also repeats its vibrational motion over a certain time interval

  time for a particle to complete a full oscillation cycle is the **period** (T)
- number of oscillations for a particle per unit time is the **frequency** (f)

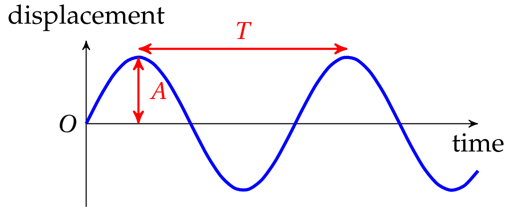
  frequency can also be defined as the number of crests passing a given point per unit time

  frequency of a wave is related to its period by $\boxed{f = \frac{1}{T}}$

  unit of frequency: $[f] = \text{Hz}$ (hertz), where $1\text{ Hz} = 1\text{ s}^{-1}$



wave pattern of all particles at one instant



vibration of one specific particle at all times

- wave energy is transferred along the direction of wave motion at a certain speed v

  in one period, wave moves forward by a distance of one wavelength

  so **wave speed** $\boxed{v = \frac{\lambda}{T}}$, or in terms of frequency, $\boxed{v = \lambda f}$

Example 10.1 When a wave travels on a water surface, the maximum depth of water is 21 cm and the minimum depth is 18 cm. What is the amplitude of the wave?

✍ amplitude is half the end-to-end distance: $A = \frac{1}{2}(21 - 18) = 1.5\text{ cm}$ □

Example 10.2 A wave travelling at 4.0 m s^{-1} has a wavelength of 50 cm, what is its period?

✍
$$v = \frac{\lambda}{T} \quad \Rightarrow \quad T = \frac{\lambda}{v} = \frac{0.50}{4.0} = 0.125\text{ s}$$
□

10.2 transverse & longitudinal waves

a wave can either be *transverse* or *longitudinal*, depending on the direction of its oscillation

10.2.1 transverse waves

a **transverse** wave has vibrations at right angle to its direction of energy transfer

- examples of transverse waves: wave on a string, surface wave on water, light wave, etc.
- for a transverse wave, greatest displacement in positive direction is called a *crest*, or a *peak*

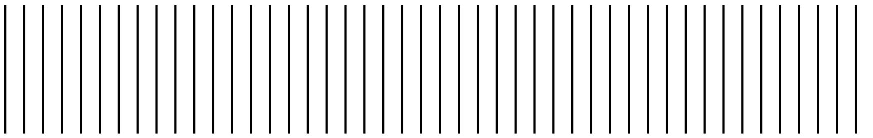
  greatest displacement in negative direction is called a *trough*

10.2.2 longitudinal waves

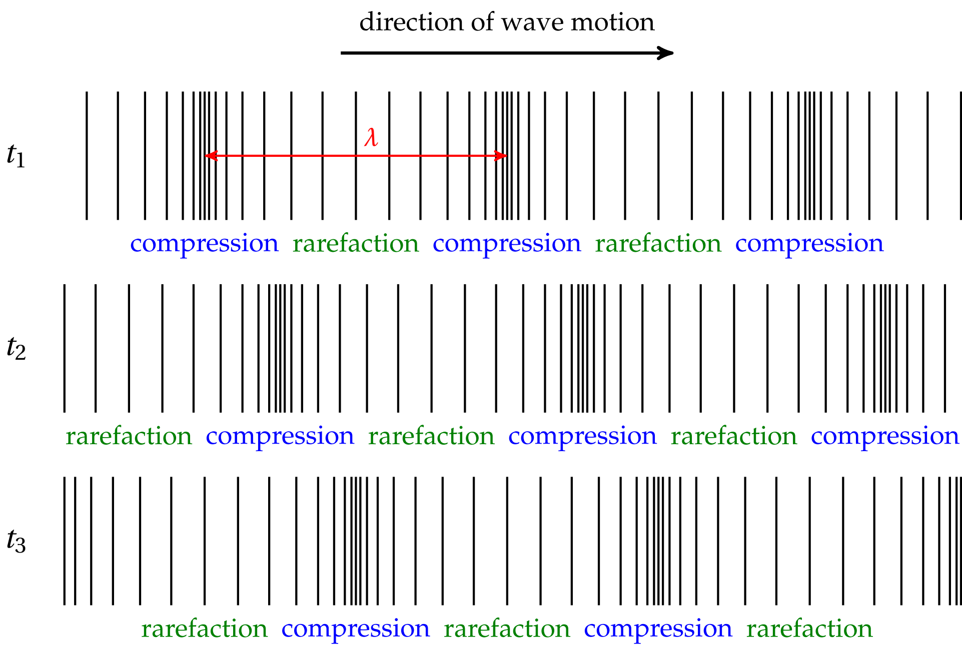
> a **longitudinal** wave has vibrations in parallel direction to energy transfer

- examples of longitudinal waves: sound waves, wave along a stretched slinky, etc.
- for a longitudinal wave, if medium gets squeezed, we say this region is a *compression*

  if a medium expands, we say this region is a *rarefaction*
- wavelength of a longitudinal wave can be defined as compression-to-compression distance



all particles at their equilibrium position before a pressure wave is set up



compression and rarefaction regions at different times ($t_1 < t_2 < t_3$)

as a longitudinal pressure wave travels in space

Example 10.3 What is the distance between a compression and a rarefaction for a sound wave of frequency 550 Hz? (Speed of sound in air is about 330 m s^{-1}.)

✍
$$d = \frac{1}{2}\lambda = \frac{v}{2f} = \frac{330}{2 \times 550} \quad \Rightarrow \quad d = 0.30 \text{ m}$$
□

10.2.3 sound waves

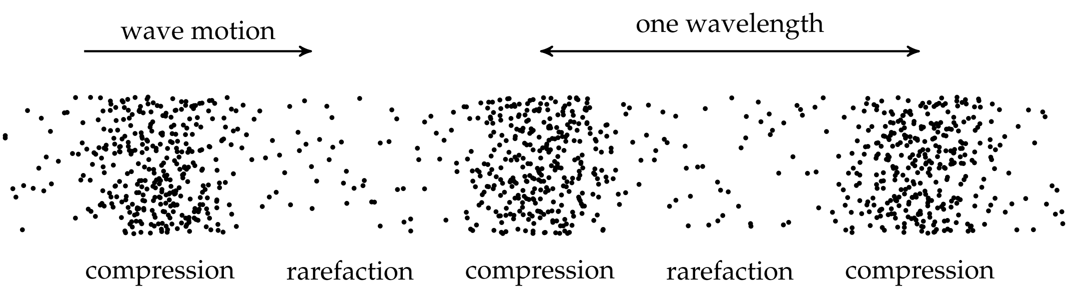
sound waves propagate via the compression and rarefaction of air (or other medium)

molecules near vibrating source is pushed away from rest positions and into their neighbours

neighbouring molecules then in turn push into their neighbours, and so on

the disturbance is transferred through the medium, forming a sound wave

➢ sound waves are longitudinal



disturbance of vibrating molecules in a sound wave

➢ propagation of sound waves require medium (air, water, steel, etc.)

sound cannot travel in vacuum

➢ speed of sound is material-dependent, but not frequency-dependent

sound in general travels faster in denser medium

- $v_{\text{air}} \approx 340 \text{ m s}^{-1}$ (under standard atmospheric pressure and room temperature)
- $v_{\text{water}} \approx 1500 \text{ m s}^{-1}$
- $v_{\text{steel}} \approx 5000 \text{ m s}^{-1}$

➢ *pitch* of a note is related to frequency of sound wave

rapid vibrations of sound source at high frequencies produce a high pitch

➢ *loudness* of sound mainly depends on amplitude of vibration

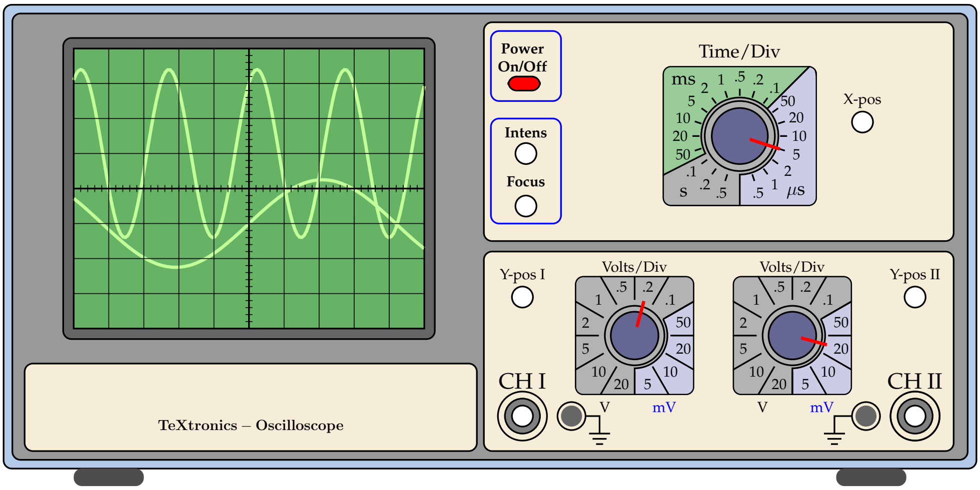
a greater amplitude means the wave is more energetic so it sounds louder

measurement of sound waves

two key apparatuses for sound measurement are the **microphone** and the **oscilloscope**

sound waves can be captured by a *microphone*, which converts sound into electrical signals electrical signals can then be sent into an *oscilloscope* (see figure[52]) for measurement

oscilloscope can be thought as an upgraded voltmeter showing how voltage varies with time



the display and controls of a typical cathode-ray oscilloscope

➢ oscilloscope displays the variation of voltage (y-axis) with time (x-axis)

➢ horizontal scale (time axis) is given by *time-base* settings

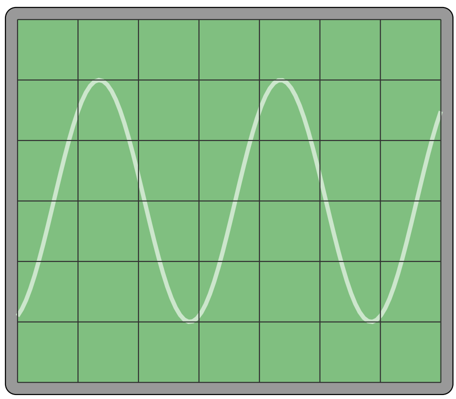
vertical scale (voltage axis) is given by *voltage gain*, or *Y-sensitivity* settings

➢ period T of the sound wave can be found using time-base settings

then frequency of sound wave is calculated: $f = \dfrac{1}{T}$

➢ voltage amplitude can be found using the voltage gain

Example 10.4 A sound wave is detected by a microphone and the trace is displayed on an oscilloscope as shown. If the time-base is set at 0.5 ms div^{-1} and the voltage gain is set at 2 V div^{-1}. What is the frequency and the amplitude of the signal?



✍ period: $T = 3 \times 0.5 \text{ ms} = 1.5 \times 10^{-3}$ s

frequency: $f = \dfrac{1}{T} = \dfrac{1}{1.5 \times 10^{-3}\text{ s}} \approx 667$ Hz

amplitude: $A = 2 \times 2 \text{ V} = 4.0$ V □

[52]The beautiful figure of the oscilloscope was created by *Hugues Vermeiren*, who generously shared the source codes on TeXample: http://www.texample.net/tikz/examples/textronics-oscilloscope/

10.3 electromagnetic waves

a wave can be either *mechanical* or *electromagnetic*, depending on whether it requires a medium

➢ waves that require a *medium* to travel are called **mechanical waves**

examples are sound waves, water waves, wave on a string, etc.

mechanical waves involve vibration of matter particles

➢ no medium is needed for propagation of **electromagnetic waves** i.e., they can travel in vacuum

10.3.1 properties of electromagnetic waves

➢ electromagnetic waves can travel in free space

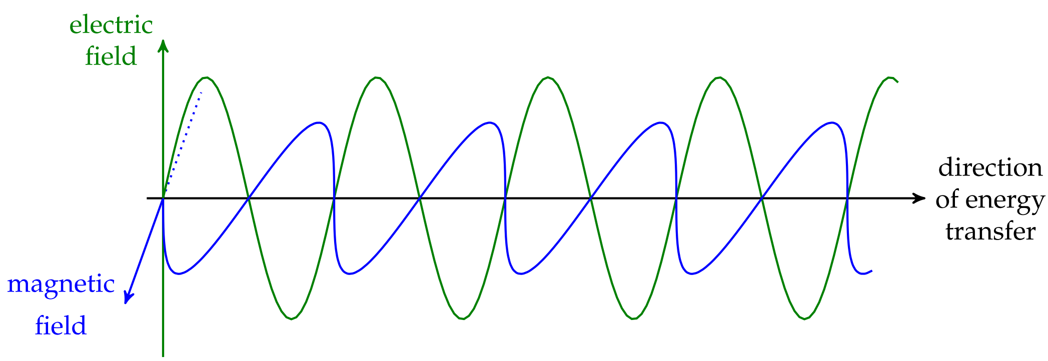
➢ electromagnetic waves involve vibrations of electric and magnetic fields

an altering electric field can generate an altering magnetic field, which then further produces a new altering electric field, so on and so forth

electric and magnetic fields then permeate through space, transferring energy and information

➢ electromagnetic waves are all transverse

vibration of electric fields and magnetic fields are both perpendicular to wave motion



variation in electric and magnetic fields for an electromagnetic wave

➢ all electromagnetic waves travel at a constant speed $c = 3.0 \times 10^8$ m s^{-1} in free space

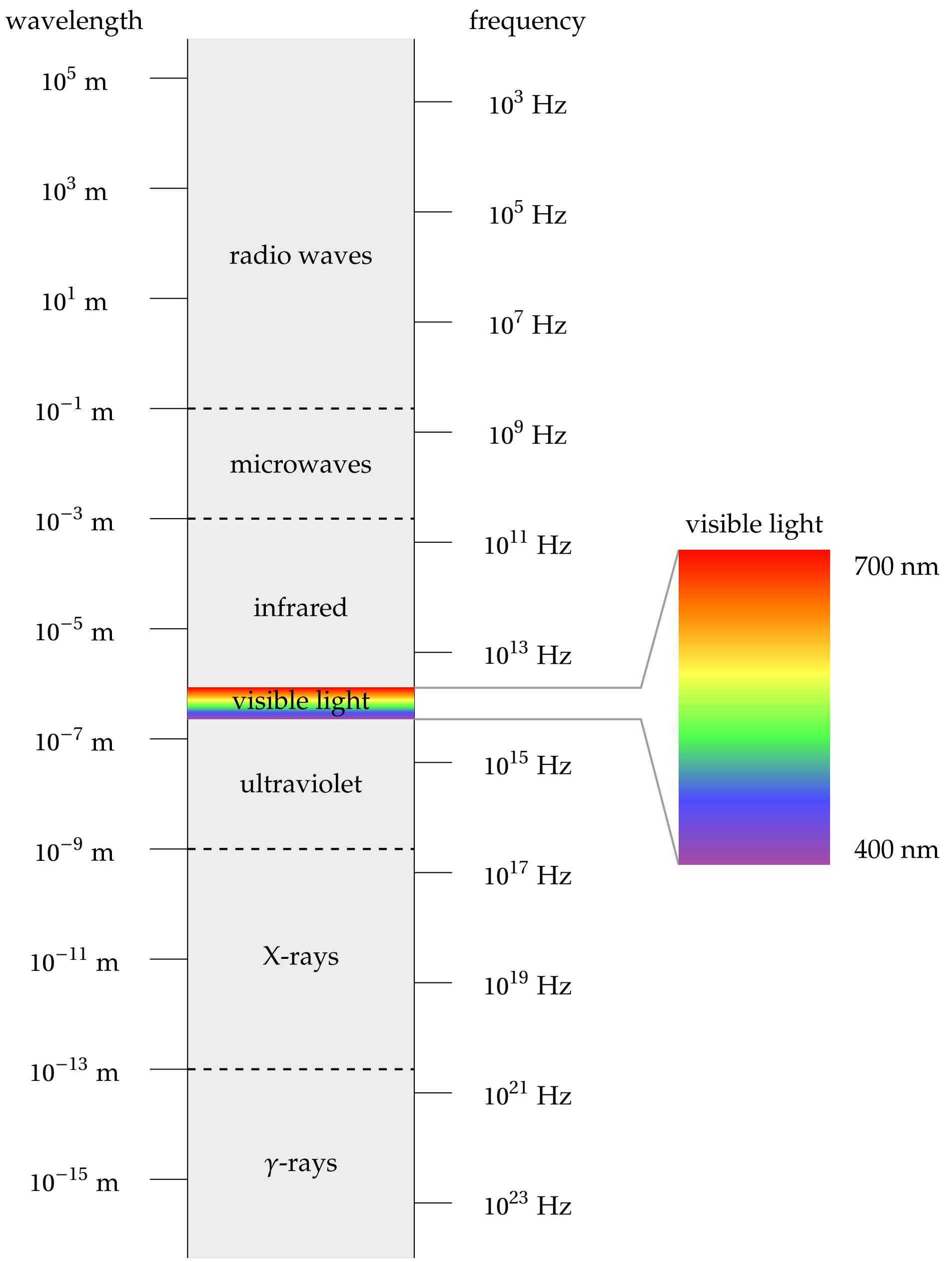
i.e., speed of light in vacuum is constant[53]

[53]The speed of light in vacuum is actually a *universal* physical constant. According to Einstein's special relativity, this is the upper limit for the speed at which matter and information can travel. This speed is also independent of the inertial reference frame one chooses, i.e., the speed of light in vacuum is the same for all observers, regardless of the motion of the source or the observer.

10.3.2 electromagnetic spectrum

electromagnetic waves come in a wide range of wavelengths and frequencies

distribution of electromagnetic radiation according to wavelength or frequency is the **electromagnetic spectrum** (see diagram)



the electromagnetic spectrum

➢ the table below shows electromagnetic spectrum in somewhat more precise details [54]

type of radiation	range of wavelength	range of frequency
radio waves	$10^{-1} \sim 10^{6}$ m	$10^{2} \sim 10^{9}$ Hz
microwaves	$10^{-3} \sim 10^{-1}$ m	$10^{9} \sim 10^{11}$ Hz
infra-red	$7\times10^{-7} \sim 10^{-3}$ m	$10^{11} \sim 10^{14}$ Hz
visible light	$4\times10^{-7} \sim 7\times10^{-7}$ m	$10^{14} \sim 10^{15}$ Hz
ultraviolet	$10^{-9} \sim 4\times10^{-7}$ m	$10^{15} \sim 10^{17}$ Hz
X-rays	$10^{-13} \sim 10^{-9}$ m	$10^{17} \sim 10^{19}$ Hz
γ-rays	$10^{-16} \sim 10^{-11}$ m	$10^{19} \sim 10^{24}$ Hz

➢ each type of electromagnetic waves has important applications in some area [55]

– radio waves

telecommunication (TV/radio broadcast, satellite communication) [56]

– microwaves

telecommunication (mobile phones, WiFi, Bluetooth, satellite communication) [57]

heating food (microwave ovens) [58]

[54]There are no precise accepted boundaries between different ranges in the electromagnetic spectrum. The boundaries are actually somewhat ambiguous. The ranges of different portions tend to overlap. For example, a large portion of the ranges of X-rays overlap with that of γ-rays, and microwaves are considered by many people as a subdivision of radio waves. Therefore, the values given here are merely supposed to give you some rough idea about the order of magnitudes for electromagnetic wavelengths and frequencies. So the point I want to make here is: do not take the borderlines too seriously.

[55]The entries listed here only include a teeny-weeny part of the uses of electromagnetic radiation, somewhat based on my personal taste. I also included a handful of explanations for the examples that I found interesting (otherwise I would not choose them), as you will see a huge load of footnotes in the next few pages. You are encouraged to do some researches as well, I can guarantee that you will not be disappointed.

[56]Having the longest wavelengths of all radiation, radio waves have the best ability to diffract around obstacles in city buildings and mountains, therefore a large area can be covered.

[57]Microwaves do not diffract sufficiently as radio waves, but they can transmit more information per unit time because they have higher frequencies. Microwaves are used in short-range telecommunication, including mobile phones, wireless networks, and bluetooth connections.

[58]Frequency of microwaves are close to the natural frequencies of the rotational motion of water molecules. When food is exposed to microwaves, the water molecules in the food resonate and vibrate

– infrared (IR)

  IR thermography (temperature monitors, thermographic cameras) [59]

  night-vision devices [60]

  IR heating (IR heat lamps, IR saunas)

  IR data transmission (remote controls, optical-fibre communication)

– visible light

  human vision [61]

– ultraviolet (UV)

  UV sterilising (drinking water treatment, disinfection of medical facilities, etc.) [62]

  fluorescent dyes (black light fluorescent paint, UV watermarks) [63]

– X-rays

  medical imaging (X-ray imaging, CT scans) [64]

  security check (luggage scanners)

  X-ray crystallography [65]

---

more violently, causing a rise in the food's temperature.

[59]All objects emit electromagnetic radiation based on their temperatures. According to the law of black-body radiation, objects near room temperature emit thermal energy as infrared radiation, so variations in the temperature can be detected.

[60]Night-vision devices convert photons (just think of them as particles of light for now) into electrons, which are amplified by a chemical and electrical process and then converted back into visible light. Infrared sources can be used to augment the available light, increasing the visibility in the dark.

[61]Human eyes are only sensitive to a small fraction of the electromagnetic spectrum. The wide variety of colours that we see is actually built up from the relative intensities of red, green and blue light collected by the three colour detectors in our eyes.

[62]Short-wavelength UV light can damage the DNA's in living organisms. A microorganism exposed to germicidal UV light might not be able to reproduce, and becomes harmless. For the same reason, overexposure to UV radiation present in the sunlight can cause sunburn, or even skin cancer.

[63]UV radiation can cause many substances to glow through chemical reactions. UV watermarks that are visible under UV light are used to prevent counterfeiting of currency, or forgery of important documents such as passports and ID cards.

[64]X-rays are very energetic and thus very penetrating. They can pass through human body easily to form an image giving information about the structures of tissues and bones.

[65]X-rays can be diffracted by the lattice of atoms in a crystal. The diffraction pattern gives information about the structure of the lattice. X-ray crystallography is a very important experimental technique to study

- γ-rays

  radiation therapies (cancer treatment) [66]

  medical imaging (PET scans) [67]

Example 10.5 A beam of electromagnetic radiation is known to have a frequency of 25 THz in vacuum. What is its wavelength?

✍ $$\lambda = \frac{c}{f} = \frac{3.00\times 10^8}{25\times 10^12} \quad\Rightarrow\quad \lambda = 1.2\times 10^{-5}\ \mathrm{m} \quad \text{(infra-red)}$$ □

10.4 wave intensity

energy can be transmitted along a wave

the degree to which energy is concentrated is called the intensity of the wave

here concentration has two meanings – concentration in time and concentration in area S[68]

intensity of a wave is defined as the power P per unit area on a cross section S: $\boxed{I = \frac{P}{S}}$

- unit of wave intensity: $[I] =$ W m^{-2}
- intensity is proportional to square of its amplitude: $\boxed{I \propto A^2}$.
- intensity of a wave decreases as it spreads out in space

  intensity at distance of r from a point source is inversely proportional to r^2: $\boxed{I \propto \frac{1}{r^2}}$ [69]

Example 10.6 If the amplitude of an incoming wave is increased by 50%, what is the increase in the wave intensity?

✍ $I \propto A^2 \quad\Rightarrow\quad \frac{I'}{I} = \left(\frac{A'}{A}\right)^2 = (1+50\%)^2 = 2.25 \quad\Rightarrow$ so intensity is increased by 125% □

the microscopic structures of new materials.

[66] γ-rays have extremely high frequencies. They are even more energetic and penetrating. They can be used to damage the DNA of cancerous tissue, and hence kill the cancerous cells as a treatment.

[67] Positron emission tomography (PET) uses a radioactive tracer to produce γ-rays within the tissues of interest. The energy and location of these γ-rays can be detected and sent to a computer to build up a 3D image of the body part.

[68] To avoid confusion, I reserved letter 'A' for wave amplitude and chose 'S' to represent an area.

[69] As a wave produced from a point source travels out by a distance r away from the source, the energy it carries is spread uniformly over the surface area of the sphere of radius r, that is: $S = 4\pi r^2$. So the intensity of a wave obeys an inverse square law: $I \propto \frac{1}{r^2}$.

Example 10.7 Two observers A and B are at a distance r_A and r_B from a point source where $r_A = 2r_B$. (a) Find the ratio of their intensities $\frac{I_A}{I_B}$. (b) Find the ratio of their amplitudes $\frac{A_A}{B_B}$.

✍

$$I \propto \frac{1}{r^2} \quad \Rightarrow \quad \frac{I_A}{I_B} = \left(\frac{r_B}{r_A}\right)^2 = \left(\frac{1}{2}\right)^2 \quad \Rightarrow \quad \frac{I_A}{I_B} = \frac{1}{4}$$

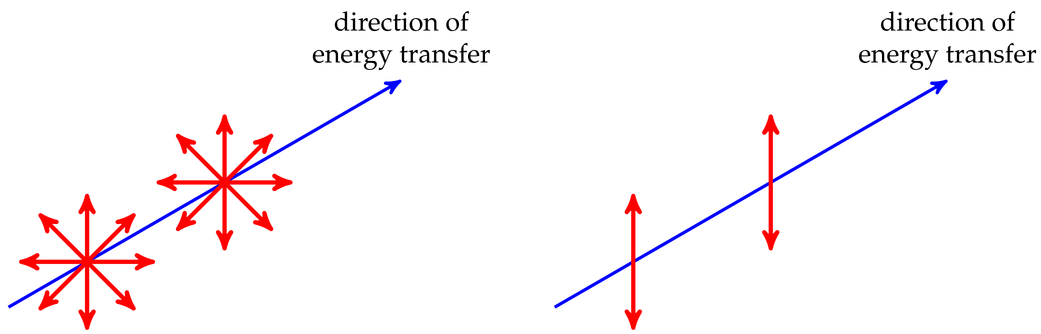
$$I \propto A^2 \quad \Rightarrow \quad \frac{A_A}{A_B} = \sqrt{\frac{I_A}{I_B}} = \sqrt{\frac{1}{4}} \quad \Rightarrow \quad \frac{A_A}{A_B} = \frac{1}{2}$$

□

10.5 polarisation

10.5.1 plane polarisation

a wave is **plane-polarised**, or simply **polarised**, if the vibration is in one single direction at right angle to the direction of propagation of energy



an unpolarised wave (wave vibrations in multiple directions) | a plane-polarised wave (wave vibrations in one direction only)

➢ polarisation[70] is a phenomenon associated with *transverse* waves only
longitudinal waves (such as sound waves) cannot be polarised
this is because longitudinal vibrations are always parallel to direction of energy transfer

➢ most important examples are polarisation of *electromagnetic waves*
regarding vibration of electromagnetic waves, we usually focus on variations of electric field

➢ sunlight, light emitted from incandescent bulbs are unpolarised
they are emitted randomly from their sources, so they contain all planes of polarisation

[70] Apart from plane polarisation where the vibrations are fixed in a single direction, there are also *circular* or *elliptical* polarisation, for which the direction of vibrations *rotate* in a plane as the wave travels.

- a few of the great many applications of polarisation are:
  - polarising sunglasses & photography [71]
  - polarised 3D films[72]
  - liquid-crystal display (LCD) technology[73]
  - radio transmission and reception[74]

10.5.2 polarisers & Malus's law

a plane-polarised light can be produced from unpolarised light using a **polariser**

a polariser blocks vibrations in all planes except the plane of polarisation[75]

direction along which vibration is allowed to pass is called the axis of the polariser

- polarised light can be sent through a second polarising filter (sometimes called an *analyser*)
- if polarised light of initial intensity I_0 has an angle θ to the axis of analyser

Malus's law states that transmitted intensity is given by: $\boxed{I = I_0 \cos^2\theta}$

this is because only electric field parallel to axis of analyser is transmitted

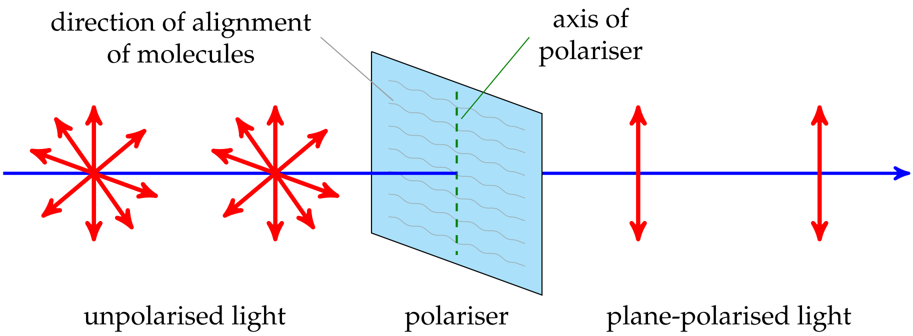
[71]When unpolarised light is reflected at the boundary of two media, the reflected beam would become partially polarised. Polarising sunglasses use this effect to reduce the sunlight reflected from shiny surfaces, such as a lake (for sightseeing tourists) or road surfaces (for drivers). For the same reason, photographers often place a polarising filter in front of the camera lens to darken skies and increase the contrast.

[72]The eyeglasses worn by viewers contain a pair of polarising filters with mutually perpendicular axes. When two images are projected onto the same screen, each filter restricts the light reaching each eye by passing only one of the images that is polarised in the same direction. This creates an 3D illusion.

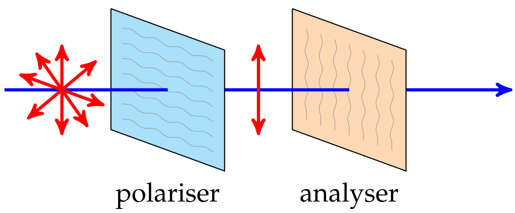
[73]Liquid crystal arrays can be realigned by applying an electric field. This causes the axis of polarization to rotate, hence backlight may be allowed to pass or be blocked, forming dark patterns of the display.

[74]Since electric currents flow in certain directions only, so the radio waves and microwaves produced from aerials are intrinsically polarised. This means the reception of signals would be sensitive to a particular direction, but totally insensitive to the normal direction. This explains why altering the orientation of antennas could greatly enhance the quality of reception.

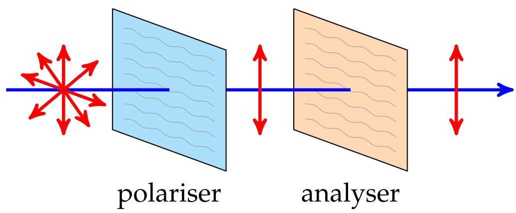
[75]What we are discussing here is just one type of polariser, often called a *linear absorptive polariser*, which is basically a synthetic transparent plastic sheet made of certain crystals. The thin sheet contains long-chain organic molecules aligned parallel to each other. When an unpolarised light passes through the polariser, there is a strong absorption of the electric field parallel to the alignment of molecules, so vibration of the electric field is allowed to pass in one direction only.



unpolarised light goes through a polariser and becomes plane-polarised



no transmission of light if polariser and analyser are aligned at right angles



polarised light is unaffected if polariser and analyser are aligned in parallel

so transmitted amplitude[76] satisfies: $A = A_0 \cos\theta$, where A_0 is the initial amplitude

recall that intensity is proportional to square of amplitude, so $I = I_0 \cos^2\theta$

Example 10.8 A beam of light polarised in the vertical direction has an amplitude A and intensity I. It passes through a polarising filter whose axis of polarisation is at 45° to the vertical. (a) What is the amplitude and the direction of polarisation of the emerging beam? (b) If the emerging beam then enters another filter whose axis of polarisation is at 75° to the vertical, what is the amplitude and intensity of the emerging beam?

✍ through first filter: $A_1 = A\cos\theta_1 = A\cos 45^\circ \quad \Rightarrow \quad A_1 = \frac{1}{\sqrt{2}}A$

light is polarised in a direction at 60° to the vertical

through second filter: $A_2 = A_1 \cos\theta_2 = \frac{1}{\sqrt{2}}A \times \cos(75^\circ - 45^\circ) \quad \Rightarrow \quad A_2 = \frac{\sqrt{6}}{4}A$

$$I_2 = \left(\frac{\sqrt{6}}{4}\right)^2 I \quad \Rightarrow \quad I_2 = \frac{3}{8}I$$

or, $I_2 = I\cos^2\theta_1 \cos^2\theta_2^2 = I \times \cos^2 45^\circ \times \cos^2(75^\circ - 45^\circ) \quad \Rightarrow \quad I_2 = \frac{3}{8}I$ □

[76]This is actually the amplitude of the electric field strength.

10.6 diffraction

a wave has the ability to bend around obstacles or pass through narrow gaps

> when a wave passes through a aperture/slit/gap/hole or encounters an obstacle, it spreads out/around the corners, this is known as **diffraction** of waves

- diffraction is a general property of all waves

  some examples of diffraction are

  - sound waves can diffract through an open door

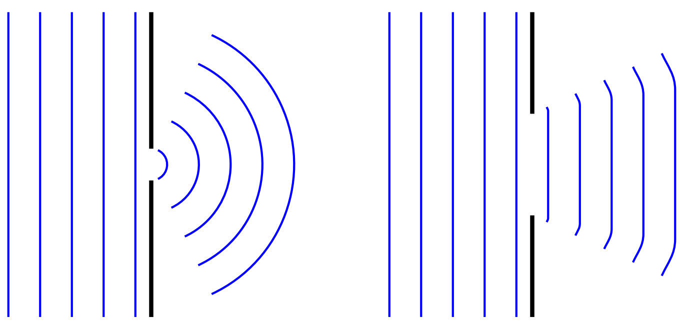
    so you can hear people in the next room talking, even though you cannot see them

  - light ray can bend when it goes through/around a small hole/small particles

    sky appears red at sunset as red light is diffracted most by dust particles in atmosphere

- a wave is diffracted most when its wavelength is close to size of aperture/obstacle

  a wave with longer wavelength usually diffracts more than a wave of shorter wavelength



diffraction of a wave as it passes through an aperture of width that is

(a) close to the wavelength, (b) greater than the wavelength

10.7 Doppler effect

> relative motion between wave source and the observer causes a change in observed frequency, this is known as the **Doppler effect**

- any type of wave can exhibit Doppler effect
- when wave source moves towards observer, a higher frequency is observed

  when wave source moves away from observer, a lower frequency is observed

➢ Doppler effect also occurs if source is at rest but observer is in motion

as long as there is *radial* motion between source and observer, there is shift in frequency

➢ Doppler effect finds its use in many areas, some examples are

- *Doppler radars*: used to measure velocity of moving target (speeding cars, tennis balls, etc.)
- *Doppler ultrasonography*: used to image blood flow in human bodies
- *astronomy*: used to study motion of stars and galaxies

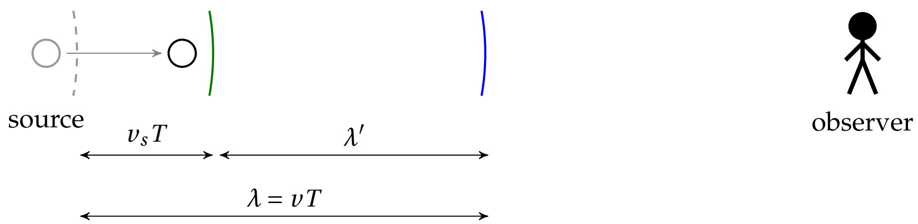
explanation for Doppler effect

suppose wave source is moving towards the observer at speed v_s

at time $t = 0$, the source emits a wavefront which travels at speed v

after one period, wavefront travels forward by a distance of vT

at same time, source moves forward by a distance of v_sT and emits a new wavefront



if source is at rest, observer simply perceives a wavelength $\lambda = vT$

now source is moving closer, a shorter wavelength λ' is perceived

since frequency is inversely proportional to wavelength, so higher frequency f' is observed

similar discussion follow for the case where source moves away from observer

➢ change in observed frequency is due to change in wavelength caused by relative motion

if source moves towards/away from observer, apparent wavelength becomes shorter/longer

Doppler effect equation

we are now ready to derive an equation for the shift of observed frequency

quantitatively, we can write: $\lambda' = \lambda \mp v_sT$ ("−"/"+" if source is moving closer/away) using $v = \lambda f$ and $f = \dfrac{1}{T}$, this becomes: $\dfrac{v}{f'} = \dfrac{v}{f} \mp \dfrac{v_s}{f}$

rearranging, one finds observed frequency is given by: $\boxed{f' = f\dfrac{v}{v \mp v_s}}$

Example 10.9 A police car moving towards you at 16 m s^{-1} sirens at 500 Hz. Given that the speed of sound in air is 340 m s^{-1}, at what frequency do you hear the siren?

✍ $$f' = f\frac{v}{v - v_s} = 500 \times \frac{340}{340 - 16} \quad \Rightarrow \quad f' \approx 525 \text{ Hz}$$ □

Example 10.10 A star emits an H_α line of wavelength 656 nm. (a) What is the frequency of this H_α line? (b) An observer on earth detects a wavelength of 680 nm, what can we say about motion of the star? (c) Find the relative speed of the star with respect to the earth.

✍ original frequency of H_α: $f = \dfrac{c}{\lambda} = \dfrac{3.00 \times 10^8}{656 \times 10^{-9}} \approx 4.57 \times 10^{14}$ Hz

observed wavelength is longer (redshift), this means the star is moving away, or receding

to find speed of star, we can consider observed frequency: $f' = \dfrac{c}{\lambda'} = f\dfrac{c}{c + v_s}$

$$\frac{3.00 \times 10^8}{680 \times 10^{-9}} = 4.57 \times 10^{14} \times \frac{3.00 \times 10^8}{3.00 \times 10^8 + v_s} \quad \Rightarrow \quad v_s \approx 1.10 \times 10^7 \text{ m s}^{-1}$$

alternatively, we can consider observed wavelength: $\lambda' = \lambda + v_s T$, or: $\Delta\lambda = \lambda' - \lambda = \dfrac{v_s}{f}$

$$(680 - 656) \times 10^{-9} = \frac{v_s}{4.57 \times 10^{14}} \quad \Rightarrow \quad v_s \approx 1.10 \times 10^7 \text{ m s}^{-1}$$ □

Chapter 11
Superposition of Waves

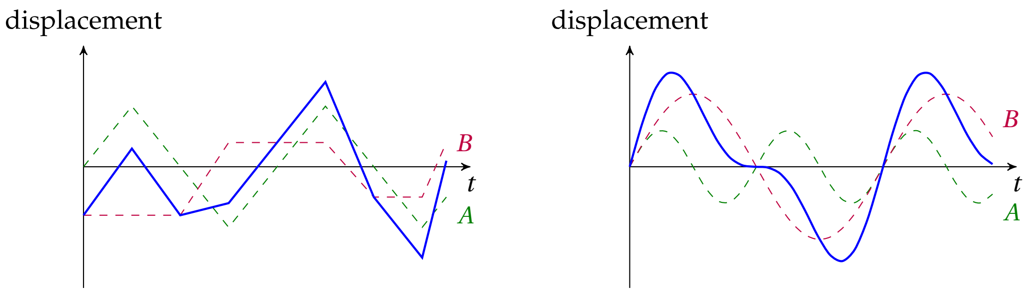
When two of more waves meet together, the resultant motion is a combination of the individuals. They can form a *resultant wave* when they overlap, after which they cross one another and neither is affected. In this chapter, we will study the principle of superposition, and look at two important consequences: the phenomenon of *interference* and *stationary waves*.

11.1 superposition of waves

11.1.1 principle of superposition

when two (or more) waves meet to form a resultant wave, the **principle of superposition** states that the resultant displacement is the vector sum of each individual displacement

Example 11.1 Displacement–time graphs for two waves A and B meeting at some point, together with the resultant wave formed at that point.



- two important situations are constructive superposition and destructive superposition
- **constructive superposition** occurs when resultant wave has greatest possible amplitude
  this happens when peaks of the two waves meet together (also trough meets trough)[77]

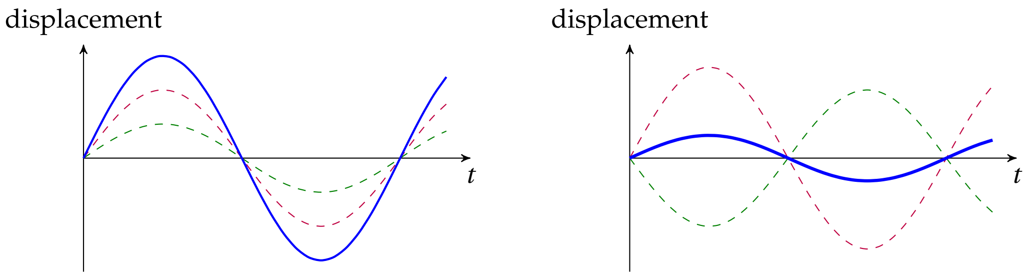
[77]This statement is implicitly for transverse waves. For longitudinal waves to superpose constructively, the regions of compression (or rarefaction) must overlap.

peaks of both waves must arrive with a time difference $\Delta t = 0,\, T,\, 2T, \cdots$

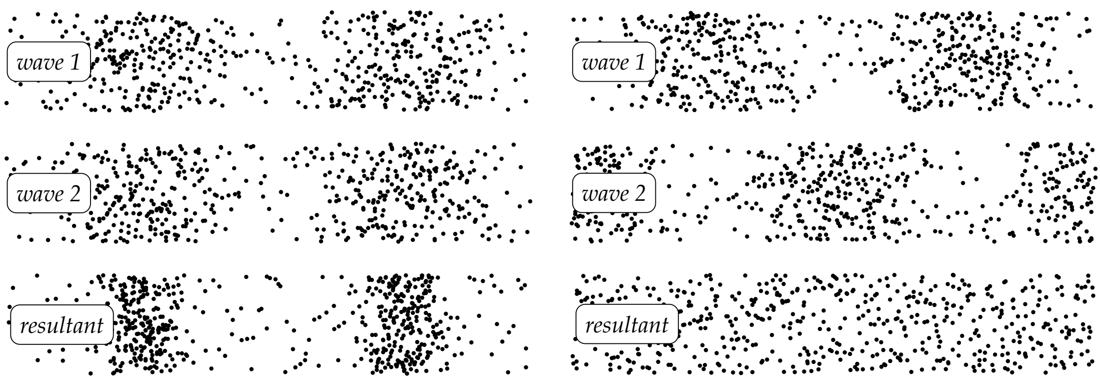
- **destructive superposition** occurs when amplitudes of each wave cancel out one another

  this happens when peak of one wave overlaps with trough of the other wave

  peaks of both waves must arrive with a time difference $\Delta t = \frac{1}{2}T,\, \frac{3}{2}T,\, \frac{5}{2}T, \cdots$



constructive and destructive superposition of two transvers waves



constructive and destructive superposition of two sound waves

11.1.2 path difference

one way to compare how much peaks of two waves differ is using the **path difference**

when two waves meet at a point, each wave travels a distance from its source

the difference between distances travelled by the two waves is the path difference ΔL

- we can tell whether two waves superpose constructively or destructively by path difference

if $\Delta L = 0,\, \lambda,\, 2\lambda, \cdots$, then superposition is constructive

if $\Delta L = \frac{\lambda}{2},\, \frac{3\lambda}{2},\, \frac{5\lambda}{2}, \cdots$, then superposition is destructive

Example 11.2 Two loudspeakers are wired to produce identical sound signals in unison. The sound wave produced has a wavelength of 80 cm. Describe the volume you hear when you are at a distance of (a) 10 m from both speakers, (b) 10 m from one speaker and 12 m from the other?

✍ path difference for case (a): $\Delta L = 10 - 10 = 0$

so constructive superposition, resultant amplitude is large, a loud sound is heard

path difference for case (b): $\Delta L = 12 - 10 = 2.0\ \mathrm{m} \quad \Rightarrow \quad \Delta L = \frac{5}{2}\lambda$

so destructive superposition, resultant amplitude is small, sound is quiet □

11.1.3 phase difference

when two waves or two vibrating particles are compared, it is also useful to describe how much one is out of step with the other in terms of their **phase difference** ($\Delta\phi$) [78]

- phase difference is measured in radians (rad) or degrees (°)
- if two waves have a phase difference $\Delta\phi = 0, 2\pi, 4\pi, \cdots$, the two waves are said to be **in phase**[79]

  in this case, peak of one wave overlaps with peak of the other
- if two waves are not in phase, then they are said to be **out of phase**
- if $\Delta\phi = \pi, 3\pi, 5\pi, \cdots$, then two waves are completely out of phase, or **anti-phase**[80]

  in this case, peak of one wave meets the trough of the other
- we can tell whether two waves superpose constructively or destructively by phase difference

if $\Delta\phi = 0, 2\pi, 4\pi, \cdots$, then superposition is constructive

if $\Delta\phi = \pi, 3\pi, 5\pi, \cdots$, then superposition is destructive

[78]More logically, we should first define the notion of *phase angle* before talking about the *phase difference*, which is simply the difference between the phase angles of two waves.

Wave motion can generally be described by sine (or cosine) functions. Since the initial displacement of a wave is not necessarily zero, one can introduce a constant term that appears in the argument of the sine function. This term shifts the origin of the sine, and we call that the *phase angle* ϕ of the wave, measured in radians (or degrees). One can think of ϕ as a number giving the fraction in a complete oscillation.

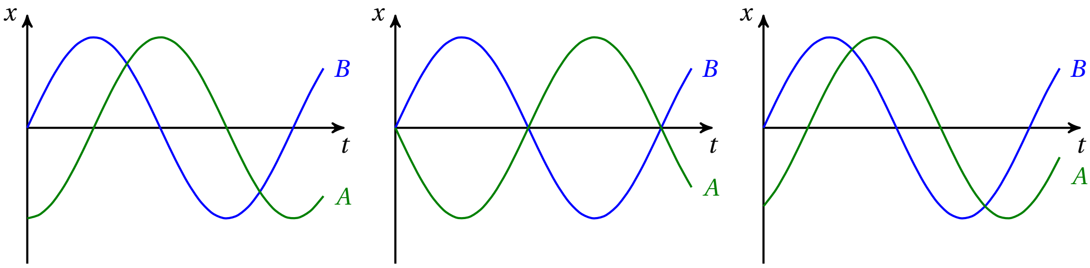
In particular, the displacement y of a transverse wave along the direction x of energy transfer can be given by: $y = A\sin\left(\frac{2\pi x}{\lambda} - \frac{2\pi t}{T} + \phi\right)$. You can verify that this indeed describes a sinusoidal wave of wavelength λ and period T travelling at speed $v = \frac{\lambda}{T}$, while ϕ determines when and where the peaks show up.

[79]If $\Delta\phi$ is given in degrees, then two waves are in phase if $\Delta\phi = 0, 360°, 720°, \cdots$

[80]If $\Delta\phi$ is given in degrees, then two waves are anti-phase if $\Delta\phi = 180°, 540°, 900°, \cdots$

➢ phase difference can be related to path difference by: $\boxed{\dfrac{\Delta\phi}{2\pi} = \dfrac{\Delta L}{\lambda} = \dfrac{\Delta t}{T}}$

Example 11.3 The diagrams below each shows the displacement of a particular point in two transverse waves. If both waves are travelling in the same direction with a wavelength of 30 cm, what is the phase difference and path difference between each wave?



(a) (b) (c)

✍ (a) $\Delta t = \frac{1}{4}T \quad \Rightarrow \quad \Delta\phi = \frac{1}{4} \times 2\pi = \frac{1}{2}\pi, \quad \Delta L = \frac{1}{4}\lambda = \frac{1}{4} \times 30 = 7.5$ cm

(b) $\Delta t = \frac{1}{2}T \quad \Rightarrow \quad \Delta\phi = \frac{1}{2} \times 2\pi = \pi, \quad \Delta L = \frac{1}{2}\lambda = \frac{1}{2} \times 30 = 15$ cm

(c) $\Delta t = \frac{1}{6}T \quad \Rightarrow \quad \Delta\phi = \frac{1}{6} \times 2\pi = \frac{1}{3}\pi, \quad \Delta L = \frac{1}{6}\lambda = \frac{1}{6} \times 30 = 5.0$ cm □

brief summary

conditions for constructive and destructive superposition can be summarised as follows:

	constructive superposition	destructive superposition
simple criteria	peak meets peak	peak meets trough
time difference	$\Delta t = n \cdot T$	$\Delta t = \left(n + \frac{1}{2}\right) \cdot T$
path difference	$\Delta L = n \cdot \lambda$	$\Delta t = \left(n + \frac{1}{2}\right) \cdot \lambda$
phase difference	$\Delta\phi = n \cdot 2\pi$ or $n \cdot 360^\circ$	$\Delta\phi = \left(n + \frac{1}{2}\right) \cdot 2\pi$ or $\left(n + \frac{1}{2}\right) \cdot 360^\circ$

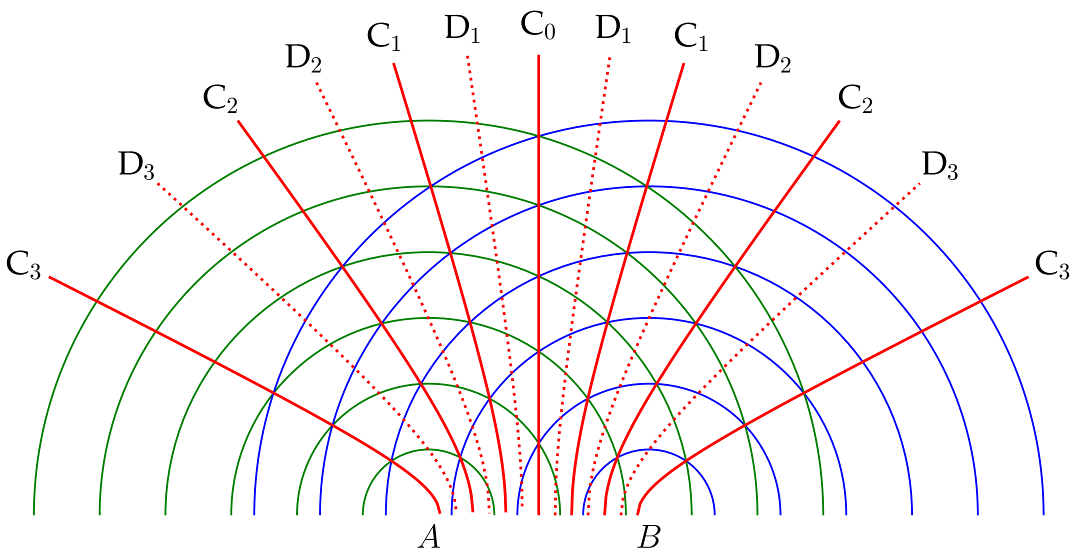
where n is a whole number: $n = 0, 1, 2, 3 \cdots$

remember that Δt, ΔL and $\Delta\phi$ are just different ways to describe how much one wave is ahead of/lag behind another wave, so they must be closely interrelated to each other

11.2 interference

11.2.1 interference

when two (or more) waves of constant phase difference meet, stable regions of constructive and destructive superposition are produced alternately, this phenomenon is called **interference**

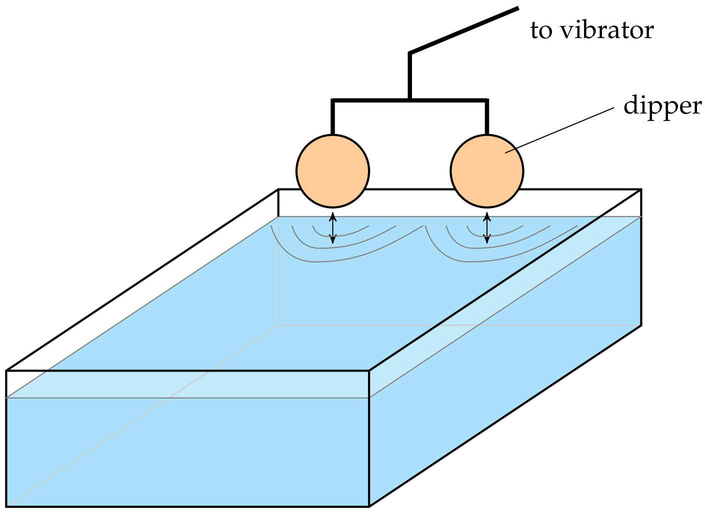


constructive (solid lines) and destructive (dashed lines) interference

of waves produced from two coherent sources A and B

- regions of constructive interference are often called *maxima*
  - points on line C_0 are of equal distance to wave sources, i.e., path difference $\Delta L = 0$
    so maxima are observed along C_0
  - points on C_1, C_2 and C_3 have path difference $\Delta L = \lambda, 2\lambda, 3\lambda$
    so maxima are also formed along C_1, C_2 and C_3
  - equivalently, points on C_n have phase difference $\Delta\phi = n \cdot 2\pi$, where $n = 0, 1, 2, 3, \cdots$
- regions of destructive interference are often called *minima*
  - points on D_1, D_2 and D_3 have path difference $\Delta L = \frac{1}{2}\lambda, \frac{3}{2}\lambda, \frac{5}{2}\lambda$
    so maxima are also formed along D_1, D_2 and D_3
  - similarly, points on D_n have phase difference $\Delta\phi = \left(n + \frac{1}{2}\right) \cdot 2\pi$, where $n = 1, 2, 3, \cdots$

- stable interference means regions of maxima/minima remain maxima/minima
  this requires **coherent** wave sources, which means constant phase difference over time
- coherent waves usually come from same source, such as
  - dippers driven by a common vibrating beam in a ripple tank
  - loudspeakers driven by the same signal generator



demonstrating interferece of water waves in a ripple tank

- interference are not observed for beams of light from different lamps or laser sources
  this is because light from different sources are usually *incoherent*[81]
  this can be overcome by dividing a single beam into several beams using a number of slits
  common apparatuses of doing so are the *double-slit* and *diffraction gratings*

Example 11.4 Two coherent waves meet in space, one of which has an amplitude of 0.30 cm and the other of 0.20 cm. How does the intensity of maxima compare with the intensity of minima?

✍ amplitude of maxima: $A_{\max} = A_1 + A_2 = 0.30 + 0.20 = 0.50$ cm

amplitude of minima: $A_{\min} = A_1 - A_2 = 0.30 - 0.20 = 0.10$ cm

ratio of intensities: $\frac{I_{\max}}{I_{\min}} = \left(\frac{A_{\max}}{A_{\min}}\right)^2 = \left(\frac{0.50}{0.10}\right)^2 \quad \Rightarrow \quad \frac{I_{\max}}{I_{\min}} = 25$ □

[81]Emission of light is associated with the electronic transitions between energy levels in an atom, which is a completely random process. So in general separate light sources are not coherent.

11.2.2 double-slit interference

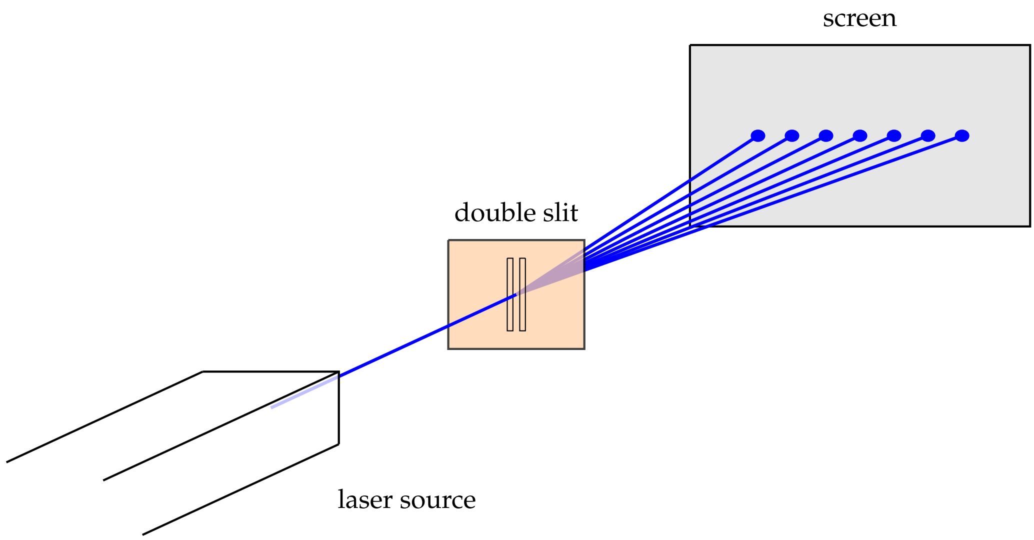
let's take a beam of light being guided through two narrow slits

light waves diffracting through the slits act as coherent wave sources

when they meet on a screen, interference pattern can be seen

since this experiment is carried out with light, alternating bright and dark fringes are formed

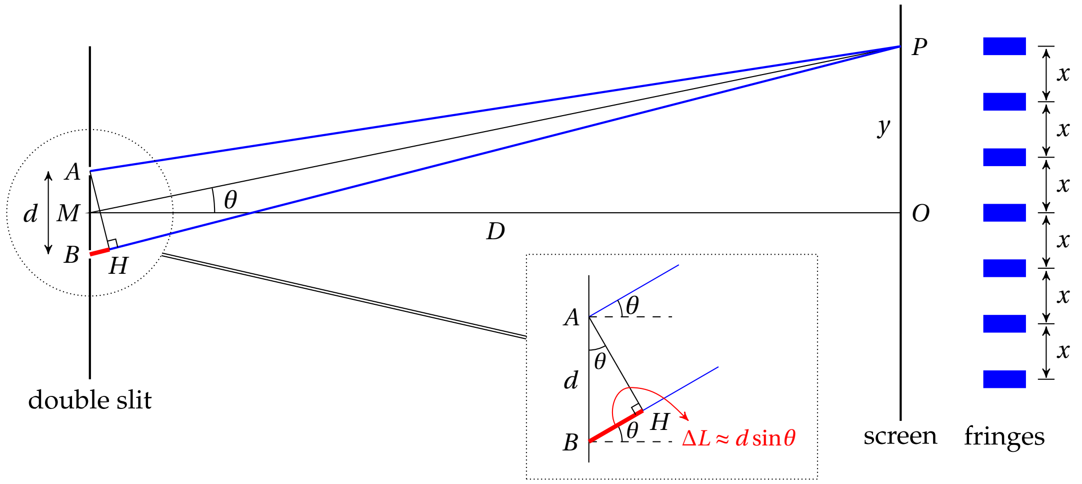
this is known as *Thomas Young*'s **double-slit experiment** [82]



Young's double-slit interference experiment

- when waves from the two slits meet with path difference $\Delta L = 0, \lambda, 2\lambda, \cdots$, maxima is formed
  these give positions where bright fringes are observed
- when waves from the two slits meet with path difference $\Delta L = \frac{1}{2}\lambda, \frac{3}{2}\lambda, \frac{5}{2}\lambda, \cdots$, minima is formed
  these give positions is where dark fringes are observed
- slit-to-screen distance is usually much larger than the slit separation, i.e., $D \gg d$
  in this case, bright fringes are nearly equally spaced with a separation of: $\boxed{x = \frac{\lambda D}{d}}$
  where d is separation of the two slits, D is slit-to-screen distance

[82] The nature of light has been argued since the history of human civilisation. It has been long debated whether light is a *wave* or it is made of *particles*. It was until the early 1800s when Thomas Young carried out the famous double-slit experiment that now bears his name, people became certain that light has wavelike properties. For a detailed review, check out the Wikipedia article: https://en.wikipedia.org/wiki/Light.



proof: consider a point P on the screen where a bright fringe is formed (see graph)

path difference between the slits must be a whole number of wavelength: $\Delta L = |PB - PA| = n\lambda$

since $D \gg d$, paths PA and PB can be considered approximately parallel

so path difference is length of the segment HB, then we have: $\Delta L \approx d \sin\theta$

in terms of θ, bright fringes are found where: $d\sin\theta = n\lambda$

to convert θ into the coordinate y, we have: $\tan\theta = \dfrac{y}{D}$

but θ is very small, small-angle approximation ($\sin\theta \approx \tan\theta$) can be applied

so positions where the bright fringes show up are: $y_n = n \times \dfrac{\lambda D}{d}$

here the coordinate subscript n labels the order of the bright fringes

this shows that the $(n+1)$-th fringe is at a fixed distance from the n-th fringe

therefore separation between neighbouring bright fringes is $x = \dfrac{\lambda D}{d}$ □

➢ altering width of slit could cause a change in *brightness* of fringes

width of slit determines intensity of emergent light passing through that slit

resultant amplitude will be affected as a consequence of the superposition principle

Example 11.5 A teacher demonstrates the double-slit experiment with a beam of red light produced from a helium-neon laser. The beam of wavelength of 632 nm is sent through a double-slit separated by 0.30 mm and passed onto a wall at about 2.0 m from the slits. (a) What is the fringe separation? (b) If the experiment is carried out using a green laser, what changes do you expect?

✍ fringe separation: $x = \dfrac{\lambda D}{d} = \dfrac{632 \times 10^{-9} \times 2.0}{0.30 \times 10^{-3}} \quad \Rightarrow \quad x \approx 4.2 \times 10^{-3}$ m

if replaced by green light, wavelength becomes shorter, so smaller fringe separation □

Example 11.6 Coherent light passes through a double slit. Initially, the light intensity from each slit is the same. State the change to the appearance of the fringes if (a) the separation between the slits is increased, (b) The width of one of the slits is reduced.

✍ (a) from $x = \frac{\lambda D}{d}$, shorter λ results in smaller fringe separation

but brightness of the fringes remains unchanged

(b) reducing width of slit causes amplitude of light from that slit to decrease

at maxima, resultant amplitude decreases, so bright fringes become less bright

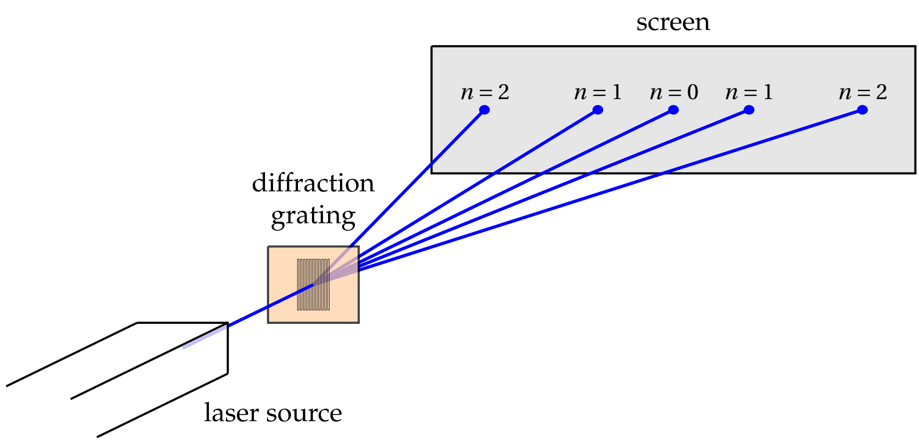
at minima, amplitudes do not cancel completely, so dark fringes become a bit brighter

but no change to separation between fringes □

11.2.3 multi-slit interference: diffraction gratings

passing light through even more slits also produces nice interference patterns

diffraction grating is such an apparatus, with hundreds or thousands of slits per mm



multi-slit interference as light passes through a diffraction grating

➢ compared to double-slit, maxima produced by diffraction grating are *sharper* and *brighter*

maxima are formed when waves from not two but many slits interfere constructively

➢ location of each maxima from a diffraction grating is given by: $\boxed{d \sin\theta = n\lambda}$

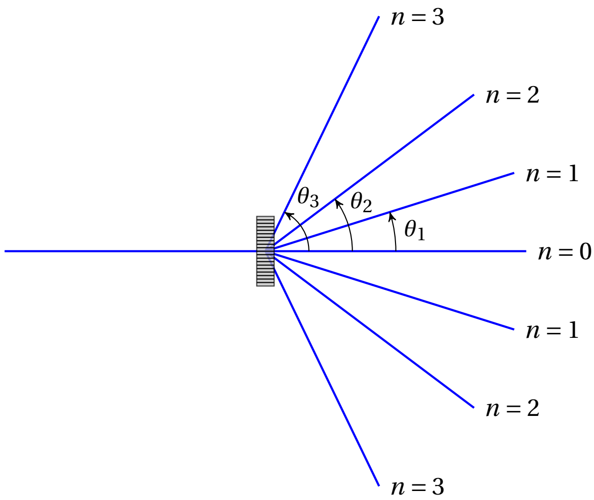
where d is the separation between adjacent slits, and light is incident *normally*

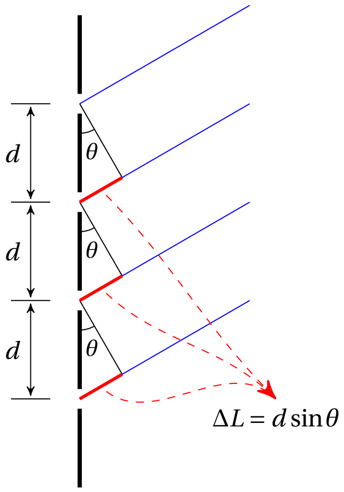
proof for constructive interference, path difference between adjacent slits must satisfy: $\Delta L = n\lambda$

since slits are so close together, light rays passing through the slits are nearly parallel

using the same trick as before, we obtain: $d \sin\theta = n\lambda$

this is known as the *diffraction grating equation*.





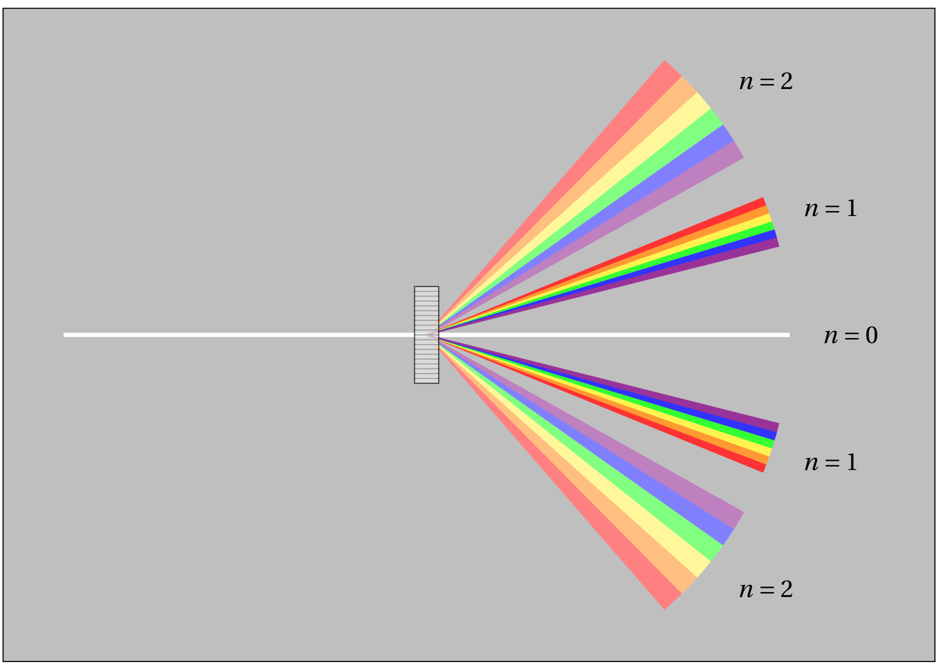
- greatest angle through which a beam of light can be diffracted is 90°

  this puts a constraint on the highest order possible:

$$n_{\max} = \frac{d\sin\theta_{\max}}{\lambda} < \frac{d\sin 90°}{\lambda} \quad \Rightarrow \quad \boxed{n_{\max} < \frac{d}{\lambda}}$$

  the value of $\frac{d}{\lambda}$ should be rounded *down* to the nearest whole number to give $n_{\max}$

- if white light is passed through diffraction gratings, *dispersion* is observed



dispersion of white light through a diffraction grating

- at zeroth order, diffraction grating equation $d\sin\theta_0 = 0\cdot\lambda$ is satisfied at $\theta_0 = 0$ for any λ
  so we observe a collection of all colours, i.e., a white central maxima
- at first order, $d\sin\theta_1 = 1\cdot\lambda$, maxima for long wavelengths appear at greater angles
  spectrum spreads into a rainbow band, with red light at outer end and violet closer in
- at higher orders, dispersion would be more spread out
  the spectra of different orders may even overlap to give complicated combinations
  we do not intend to discuss this further

➢ diffraction grating is an import device in analysis of light
it is widely used in measurement of wavelength of light

Example 11.7 Light of 632 nm wavelength produces first-order maxima at 16° when passing through a diffraction grating at right angles. How many lines per millimetre are there in the diffraction grating?

✍ slit separation: $d = \dfrac{n\lambda}{\sin\theta} = \dfrac{1\times 632\times 10^{-9}}{\sin 16^\circ} \quad\Rightarrow\quad d \approx 2.29\times 10^{-6}$ m

number of lines in 1 mm: $N = \dfrac{1 \text{ mm}}{2.29\times 10^{-6}\text{ m}} = \dfrac{1\times 10^{-3}}{2.29\times 10^{-6}} \quad\Rightarrow\quad N\approx 436$ □

Example 11.8 Green light of 510 nm wavelength is incident normally on a diffraction grating with slit separation of 2.0 μm. (a) What is the highest order seen? (b) how many maxima are produced? (c) If blue light is used for the same diffraction grating experiment, suggest the effect on the diffraction pattern.

✍ to find highest order, we have: $n_{\max} < \dfrac{d\sin 90^\circ}{\lambda} = \dfrac{2.0\times 10^{-6}}{510\times 10^{-9}} \approx 3.9 \quad\Rightarrow\quad n_{\max} = 3$

1st, 2nd, 3rd-order on either side and 0th-order at centre, so $2\times 3+1 = 7$ maxima are formed

blue light has shorter wavelength, so more maxima will be produced with blue light, and separation between each maxima would be closer □

Example 11.9 Light of wavelength 630 nm produces a second-order maxima at angle of 60° when it is directed at a diffraction grating. If visible light of another wavelength can also give a maximum at the same angle, what is this wavelength?

✍ using diffraction grating equation: $d\sin\theta = n\lambda$, we find $n\lambda$ is constant for fixed θ

from information given in the question, this constant is $2\times 630 = 1260$ nm

if $n = 1$, $\lambda = 1260$ nm, which would be infra-red

if $n = 3$, $\lambda = 420$ nm, which would be visible (violet)

if $n \geq 4$, $\lambda \leq 315$ nm, which would be ultraviolet

so the desired wavelength is 420 nm □

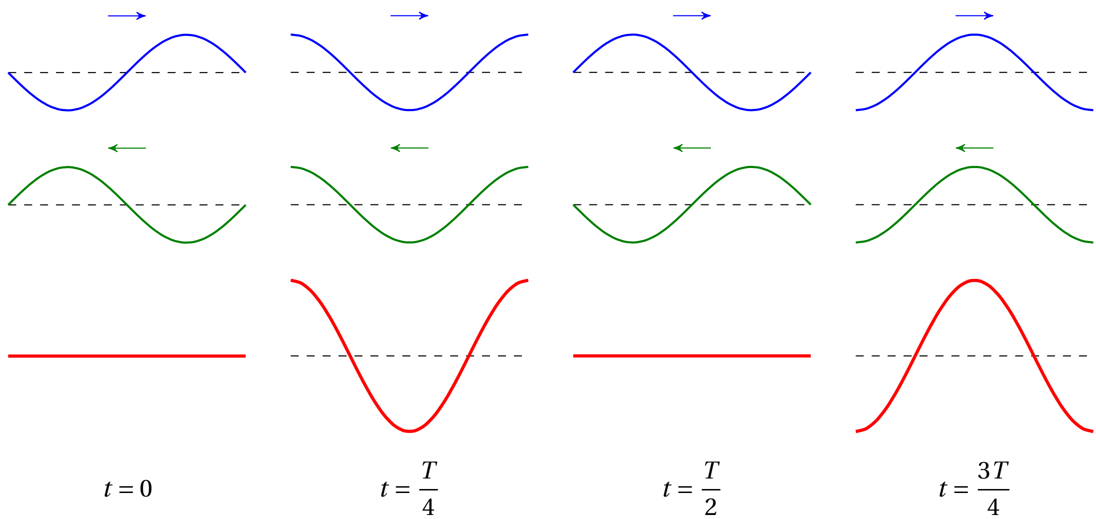
11.3 stationary waves

11.3.1 formation of stationary waves

stationary waves are often formed when a travelling wave is *reflected* at a some point

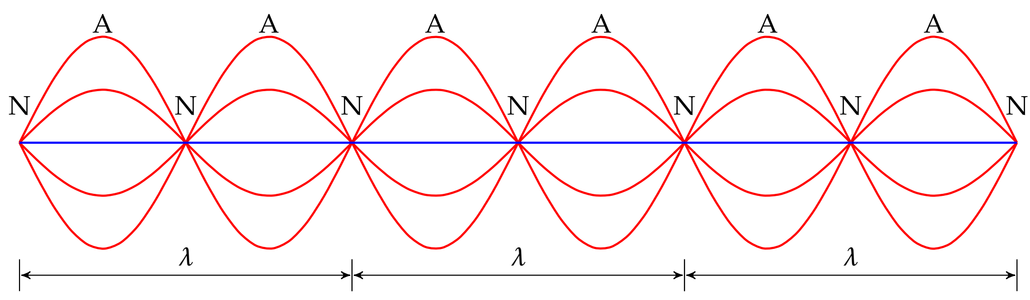
forward-moving wave and backward-moving reflected wave combine to form stationary wave

> when two waves of same frequency and same speed travel in opposite directions and meet together, they superpose to form stable regions of constructive and destructive interference, this forms a **stationary wave**, also known as a **standing wave**



formation of a stationary wave (red) due to the superposition of one wave travelling to the right (blue) and another wave travelling to the left (green)

from this one can visualise how the pattern of a stationary wave varies with time



variation of a stationary wave pattern with time

- there are points in the wave where destructive interference always occurs
  these points are called **nodes**, which have zero amplitudes
- points oscillating with the greatest amplitudes are called **anti-nodes**
- points other than nodes and anti-nodes oscillate with different amplitudes
- distance between two adjacent nodes is half of one wavelength
  wavelength can be estimated as: $\boxed{\lambda = 2 \times \text{node-to-node distance}}$

11.3.2 stationary wave pattern & boundary conditions

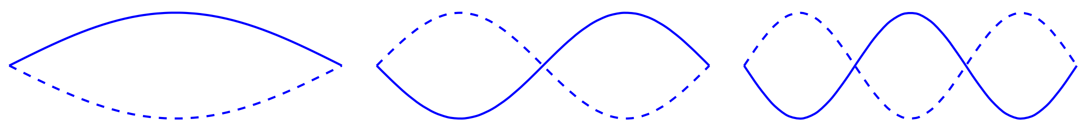
patterns of stationary waves depend heavily on the boundary conditions

- the end can be *fixed*, also called a *closed* end, to form a *node*
- the end can also be *free* to oscillate, known as an *open* end, to form *an anti-node*

stationary wave between two fixed ends

with both ends fixed, the two ends are both nodes

three patterns with the longest wavelengths and lowest frequencies are shown

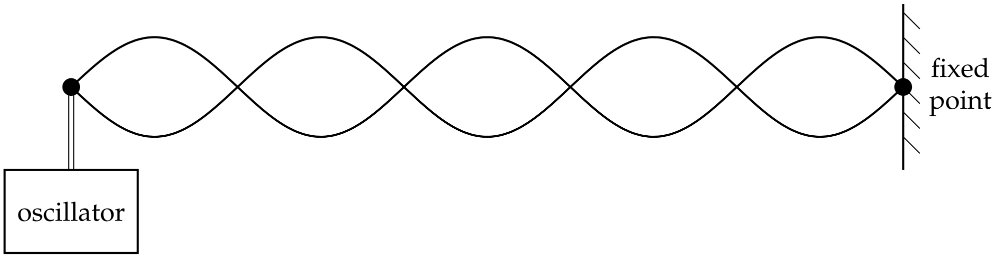


fundamental mode 2nd harmonic 3rd harmonic

- longest-wavelength mode is called the *fundamental mode*, it also has lowest frequency
  modes with higher frequencies are usually referred to as *excited modes*, or *harmonics*[83]
  in many texts different modes are labelled as 1st, 2nd, 3rd *harmonics*, etc.
- if two ends are separated by a distance of L, then wavelengths for different modes are:
  $$\lambda_1 = 2L \qquad \lambda_2 = L \qquad \lambda_3 = \frac{2L}{3} \qquad \cdots \qquad \lambda_n = \frac{2L}{n}$$
  making comparison with the fundamental wavelength λ_1 λ_1, we have
  $$\lambda_2 = \frac{\lambda_1}{2} \qquad \lambda_3 = \frac{\lambda_1}{3} \qquad \cdots \qquad \lambda_n = \frac{\lambda_1}{n}$$
- since wave speed remains constant as wave travels through the same medium
  frequency of each mode is inversely proportional to its wavelength, hence,
  $$f_2 = 2f_1 \qquad f_3 = 3f_1 \qquad \cdots \qquad f_n = nf_1$$

[83]These modes are closely related to the notes played on musical instruments.

Example 11.10 *Melde's experiment* demonstrates a stationary wave formed on a vibrating string using an oscillator set at a particular frequency. Given that the length of the string is 80 cm. (a) What is the wavelength of this wave? (b) If one wants to construct a stationary of longer wavelength, suggest what changes can be made?



Melde's experiment: demonstration of stationary waves on a tense string

✍ five loops so five half-wavelengths $\Rightarrow$ $L = \frac{5}{2}\lambda$ $\Rightarrow$ $\lambda = \frac{2}{5}L = \frac{2}{5} \times 80$ $\Rightarrow$ $\lambda = 32$ cm

to have larger wavelength, one could do either of the following:

– use a string with greater length (while keeping frequency of oscillator unchanged)

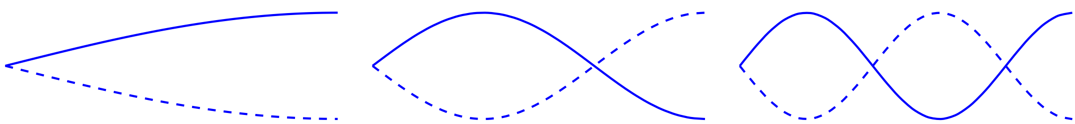
– reduce the frequency of oscillation (constant wave speed, so λ increases)

– increase tension in string (wave speed increases[84] for same frequency, so λ increases) □

stationary wave between one fixed end and one open end

in this case, we have a node at one end and an anti-node at the other end

again we give the three patterns with the longest wavelengths and lowest frequencies



fundamental mode 2nd harmonic 3rd harmonic

➢ longest-wavelength/lowest-frequency mode is also called the *fundamental mode*

excited modes are called *harmonics* as before

➢ wavelengths for allowed patterns under the boundary conditions are

$$\lambda_1 = 4L \qquad \lambda_2 = \frac{4L}{3} \qquad \lambda_3 = \frac{4L}{5} \qquad \cdots \qquad \lambda_n = \frac{4L}{2n-1}$$

[84]The speed v of a progressive wave travelling along a string of length L and mass m can be given by the equation: $v = \sqrt{\frac{TL}{m}}$, where T is the tension in the string.

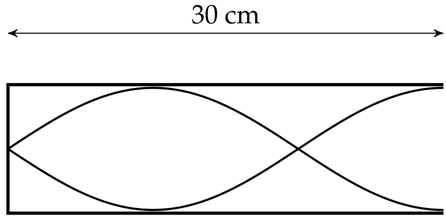
comparing with the fundamental wavelength λ_1, we find

$$\lambda_2 = \frac{\lambda_1}{3} \qquad \lambda_3 = \frac{\lambda_1}{5} \qquad \cdots \qquad \lambda_n = \frac{\lambda_1}{2n-1}$$

again the wave speed should be the same for all modes, so the allowed frequencies are

$$f_2 = 3f_1 \qquad f_3 = 5f_1 \qquad \cdots \qquad f_n = (2n-1)f_1$$

Example 11.11 The graph shows a standing sound wave produced in an air column in a closed pipe of length 30 cm. The frequency of this sound wave is 825 Hz. (a) What is the wavelength of the sound wave? (b) Calculate the speed of sound. (c) What is the frequency of the fundamental mode for sound wave in this air column? (d) Describe the motion of an air molecule near the open end.



✍ (a) wavelength of this excited mode: $\lambda = \frac{4}{3}L = \frac{4}{3} \times 30 \times 10^{-2} \quad \Rightarrow \quad \lambda = 0.40 \text{ m}$

(b) wave speed: $v = \lambda f = 0.40 \times 825 \quad \Rightarrow \quad v = 330 \text{ m s}^{-1}$

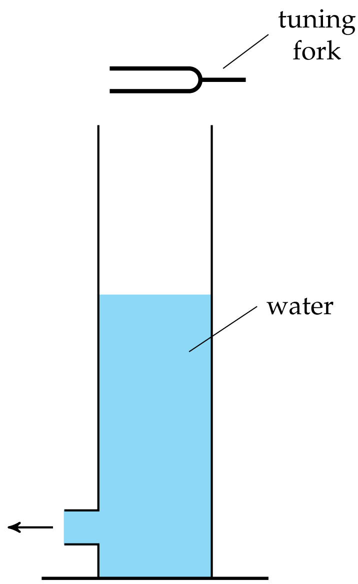
(c) wavelength of fundamental mode: $\lambda = 4L = 4 \times 30 \times 10^{-2} \quad \Rightarrow \quad \lambda = 1.2 \text{ m}$

fundamental frequency: $f = \frac{v}{\lambda} = \frac{330}{1.2} \quad \Rightarrow \quad f = 275 \text{ Hz}$

(d) note that an anti-node is formed near open end, also sound wave is longitudinal

so air molecule near open end vibrate *horizontally* with greatest amplitude □

Example 11.12 A tuning fork is held above a tall cylinder which is initially filled with water. The water level, measured from the bottom of the cylinder, is lowered at a constant rate. A loud sound is first heard when the water level is at 60.0 cm, and the next loud sound is heard when the water level is at 26.8 cm. It is known that the speed of sound in air is 340 m s^{-1}. What is the frequency of the sound wave produced by the tuning fork?



✍ loud sound is heard when stationary wave is set up

node is formed at surface of water

node-to-node distance is $60.0 - 26.8 = 33.2$ cm

wavelength of this wave: $\lambda = 2 \times 33.2 \text{ cm} = 0.664 \text{ m}$

frequency: $f = \frac{v}{\lambda} = \frac{340}{0.664} \quad \Rightarrow \quad f \approx 512 \text{ Hz}$ □

stationary wave between two open ends

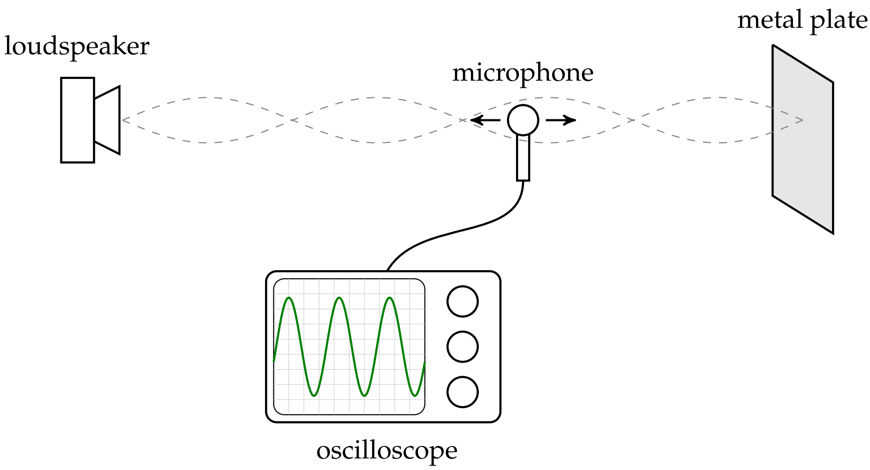
the result is left as an exercise for the reader to prove

you shall verify that the allowed wavelengths are: $\lambda = 2L, L, \frac{2}{3}L, \cdots, \frac{2L}{n}, \cdots$

frequencies of the fundamental and excited modes go as $f = f_1, 2f_1, 3f_1, \cdots, nf_1 \cdots$

11.3.3 measurement of speed of sound with stationary waves

speed of sound waves can be measured using stationary waves as follows



arrangement of apparatuses for setting up a stationary sound wave

- to set up a stationary sound wave, we need two travelling waves in opposite directions
  sound wave can be produced by a *loudspeaker* (coupled to a *signal generator*)
  this forward-travelling wave is reflected at a *metal plate*
  reflected wave superposes with wave from loudspeaker to give a stationary wave
  wave intensity can be detected by a *microphone* and observed on an *oscilloscope*
- to verify whether stationary wave is formed, we check whether nodes are formed
  positions of the metal plate and the microphone are adjusted carefully
  if there exists a location where no signal is detected, this means microphone is at a node
  now a stationary wave has been formed between loudspeaker and reflector
- to find wavelength of sound wave, we slowly move microphone to the next node
  wavelength of sound wave: $\lambda = 2 \times$ distance between neighbouring nodes

➢ to find frequency of sound wave, we move microphone to anywhere that is not a node
from time base of the oscilloscope, we can work out period T of the wave
frequency is then given by the formula: $f = \dfrac{1}{T}$

➢ finally, speed of the sound wave[85] can be calculated: $v = \lambda f$

11.3.4 stationary waves & progressive waves

there are quite a few differences between a stationary wave and a progressive wave

➢ progressive wave can transfer energy from one place to another
but for stationary waves, vibrational energy is not transferred due to existence of nodes

➢ different points in general have different amplitudes over a stationary wave
nodes have zero amplitude, anti-nodes have greatest amplitudes
other points have various amplitudes depending on their positions
but for a progressive wave, all points have the same amplitude
displacements of each point at a particular moment can be all different
but the largest displacements they can reach are the same (if no loss of energy)

➢ phase difference between any two points in a progressive wave can take any value
different points reach their peaks at different times for a progressive wave
so $\Delta\phi$ can be anything from 0 to 2π (or anything from 0 to 360°)
phase difference between any two points in a stationary wave is either 0 or π (or 180°)
for points between adjacent nodes, they all reach their highest at same time
despite differences in amplitudes, these points are all in phase, so $\Delta\phi = 0$
for points separated by one node, they are anti-phase
this is because when one is at its peak, the other is at its trough, so $\Delta\phi = \pi$ (or 180°)

[85]We should bear in mind that stationary waves do not propagate through space, so rigorously speaking, it is incorrect to say the speed of a stationary way. The wave speed calculated in this way refers to the speed of either of the two travelling waves that give rise to the stationary wave through superposition.

CHAPTER 12
Electrical Quantities & Components

12.1 electrical quantities

12.1.1 electric charges

electric charge is the property of matter that causes it to experience a force in an electric field

electric charges come in two types, *positive* charges (+ve) and *negative* charges (-ve)

you should be familiar that like charges repel each other, while opposite charges attract

interaction between electric charges plays a central role in electrical phenomena

➢ electric charges are measured in coulombs: $[Q] = \mathrm{C}$

➢ all matter is made of atoms, which consist of a nucleus (+ve) and electrons (-ve)

each electron has a charge of $-e$, where $e = 1.60 \times 10^{-19}$ C is the **elementary charge**

➢ a body becomes electrically-charged by losing or gaining electrons

- an object that gains electrons from elsewhere becomes negatively-charged
- an object that loses electrons becomes positively-charged

➢ an object can only lose or gain a whole number of electrons

charge of any object must be an integer multiple of the elementary charge: $\boxed{Q = Ne}$

we say the charge of an object is **quantised**

➢ electric charges cannot be created from nowhere nor be destroyed

charges can be transferred from one object to another, but the total amount is constant

this is a law of conservation, known as the **conservation of electric charges**

Example 12.1 A metal sphere carries a net charge of $+2.4 \times 10^{-9}$ C. Suggest whether the sphere has gained additional electrons or lost electrons, and also calculate how many electrons has the sphere gained or lost?

✍ the sphere is positively-charged, so it has lost some of its electrons

number of lost electrons: $N = \frac{Q}{e} = \frac{2.4 \times 10^{-9}}{1.60 \times 10^{-19}} \quad \Rightarrow \quad N = 1.5 \times 10^{10}$ □

12.1.2 electric current

flow of electric charges gives rise to electric currents

> **electric current** I is defined as the amount of charge flow Q per unit time: $\boxed{I = \dfrac{\Delta Q}{\Delta t}}$

- unit of electric current: $[I] = \text{A}$ (ampere)

  note that ampere is one of the S.I. base units[86][87]
- direction of a current is defined as the direction of flow of positive charges

  but current flowing in a conductor is usually due to motion of negatively-charged electrons

  so conventional current is in opposite direction to the electron flow
- to find charge flow from a current, we can use $Q = It$ if this is a constant current

  if the current varies with time, area under the I-t graph can give the charge flow

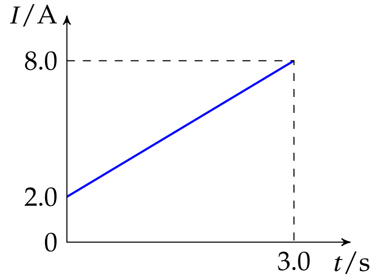
  or one can also seek to find the average current and then proceed with $Q = \bar{I}t$
- current in a circuit can be measured with an **ammeter**

Example 12.2 A current of 25 mA flows in a wire. How many electrons have passed a point in the wire in one hour?

✍ charge flow: $Q = It = 25 \times 10^{-3} \times 3600 \quad \Rightarrow \quad Q = 90\text{ C}$

number of electrons: $N = \dfrac{Q}{e} = \dfrac{90}{1.60 \times 10^{-19}} \quad \Rightarrow \quad N \approx 5.6 \times 10^{20}$ □

Example 12.3 The current in a lamp is increased uniformly from 2.0 A to 8.0 A over 3.0 s. What is the charge flow?



✍ average current during this time is: $\bar{I} = \dfrac{2.0 + 8.0}{2} = 5.0\text{ A}$

charge flow is then: $Q = \bar{I}t = 5.0 \times 3.0 = 15\text{ C}$

we can also use area under I-t graph

$$Q = \frac{1}{2} \times (2.0 + 8.0) \times 3.0 = 15\text{ C}$$ □

[86] A current of 1 A is defined as the following: when two parallel infinitely-long straight wires separated by a distance of 1 m carry two equal currents such that the force per metre between them is 2.0×10^{-7}N, then this current is of 1 A.

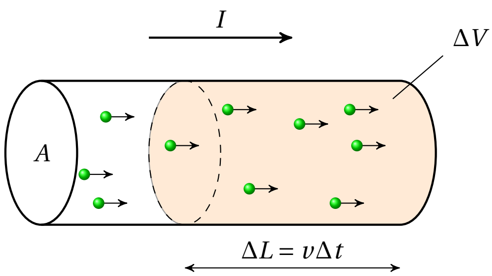
[87] It might sound strange to you that the notion of electric current is defined based on electric charge, but the unit of charge is defined based on the unit of current: 1 C = 1 A s. Sometimes physicists are like nuts.

microscopic view of electric currents

flow of electric current is essentially due to motion of charge carriers[88]

naively, current would depend on number of charge carriers and how fast they move

let's take cross section A of a material, and see what determines the electric current



let's define n being the number density of charge carriers for the material

then total number N of charge carriers in a volume V can be given by: $N = nV$

if each carrier has a charge of q and moves with a speed v, then they contribute to a current:

$$I = \frac{\Delta Q}{\Delta t} = \frac{\Delta N q}{\Delta t} = \frac{n\Delta V q}{\Delta t} = \frac{nA\Delta L q}{\Delta t} = \frac{nAv\Delta t q}{\Delta t} \quad \Rightarrow \quad \boxed{I = nAvq}$$

➢ for metals, free electrons act as charge carriers, the equation becomes: $I = nAve$

➢ charge carriers have a distribution of speeds as they flow

so v is actually an average velocity, called the *mean drift velocity*

➢ n is number density of charge carriers, whose value depends on type of material

good conductors have large n, while semiconductors have smaller n

Example 12.4 A length of silver wire of a diameter of 1.2 mm carries a current of 2.0 A. The electron number density in silver is 5.9×10^{28} m^{-3}. What is the mean drift velocity of the electrons?

✍ $$v = \frac{I}{nAe} = \frac{I}{n \cdot \frac{1}{4}\pi d^2 \cdot e} = \frac{2.0}{5.9 \times 10^{28} \times \frac{1}{4}\pi \times (1.2 \times 10^{-3})^2 \times 1.60 \times 10^{-19}} \approx 1.9 \times 10^{-4} \text{ m s}^{-1}$$

this calculation shows that each free electron actually travel at very low speeds in a wire □

Example 12.5 A metal conductor and a piece of semi-conductor of the same cross section are connected in series. When an electric current is driven through them, compare the mean drift speed of the charge carriers in the two materials.

✍ same current and same cross section, so mean drift speed $v = \frac{I}{nAe} \propto \frac{1}{n}$

metal has larger n, so free electrons move at relatively lower speeds

semi-conductor has smaller n, so charge carriers move at relatively higher speeds □

[88] In most metallic conductors, charge carriers are those *free electrons*. Charge carriers can also be positively-charged *holes* in some semi-conductors or *ions* in chemical solutions.

12.1.3 p.d. & e.m.f.

p.d. and e.m.f.[89] are both defined as the energy transfer per unit charge

both terms are often called *voltages* in less formal contexts

p.d. (potential difference) between two points is the energy needed to move a unit positive charge from one point to the other: $\boxed{V = \dfrac{W}{Q}}$

e.m.f. (electromotive force) of a power supply is the energy needed to move a unit positive charge around a complete circuit: $\boxed{\mathcal{E} = \dfrac{W}{Q}}$

➢ unit of measurement for p.d./e.m.f.: $[V] = [\mathcal{E}] = \mathrm{V}$ (volt), where $1\ \mathrm{V} = 1\ \mathrm{J\,C^{-1}}$

➢ p.d. differs from e.m.f. mainly due to the forms of energy transfer involved

p.d. is amount of electrical energy transferred *into* other forms (heat, light, mechanical, sound, etc.) per unit charge across a component (resistor, lamp, motor, etc.)

e.m.f. is amount of electrical energy transferred *from* other forms (chemical, mechanical, wind, nuclear, etc.) per unit charge in a power supply (battery, generator, etc.)

➢ p.d. can tell direction of current flow through a component

current always flows from a higher potential to a lower potential

➢ p.d. across a component can be measured with a **voltmeter**

Example 12.6 How much electrical energy is transformed into thermal energy when a charge of 10 C flows through a heater whose p.d. across is 25 V?

✍ $$V = \frac{W}{Q} \quad \Rightarrow \quad W = VQ = 25 \times 10 = 250\ \mathrm{J}$$ □

Example 12.7 A fully-charged 12 V car battery can supply a total electrical energy of 0.90 MJ. The starter motor of the car requires an average current of 150 A for a period of 2.0s. The battery is not able to recharge due to a fault. How many times can the starter motor be used?

✍ total charge that can be supplied: $Q_{\text{total}} = \dfrac{W}{\mathcal{E}} = \dfrac{0.90 \times 10^6}{12} = 75000\ \mathrm{C}$

charge needed for each start: $Q = It = 150 \times 2.0 = 300\ \mathrm{C}$

number of starts possible: $N = \dfrac{Q_{\text{total}}}{Q} = \dfrac{75000}{300} = 250$ □

[89] e.m.f. stands for *electromotive force*, which is a miserably misleading term due to historical reasons. It has nothing to do with a force. It is better to simply use the abbreviation e.m.f.

Example 12.8 Electrons beams are accelerated between a heated cathode and an anode with a potential difference of 40 V. What is the speed of the electrons when they reach the anode?

✍ electrical potential energy is transformed into kinetic energy of electrons:

$$qV = \frac{1}{2}mv^2 \quad \Rightarrow \quad v = \sqrt{\frac{2qV}{m}} = \sqrt{\frac{2 \times 1.60 \times 10^{-19} \times 40}{9.11 \times 10^{-31}}} \approx 3.75 \times 10^6 \text{ m s}^{-1} \qquad \square$$

12.1.4 resistance

> **resistance** R of an electrical component is defined as the potential difference across divided by the current flowing through it: $\boxed{R = \frac{V}{I}}$

➢ unit of resistance: $[R] = \text{V A}^{-1} = \Omega$ (ohm)

when a p.d. of 1 V is applied and current flow is 1 A, then resistance is 1 Ω

➢ resistance of a component does not depend on p.d. applied or current through it

➢ resistance R of a component mainly depends on the following factors:

- R is proportional to length L of the component: $R \propto L$
- R is inversely proportional to cross-sectional area A : $R \propto \frac{1}{A}$
- R depends on the material of the component

this can be written as: $\boxed{R = \rho \frac{L}{A}}$

the constant ρ is called the **resistivity** of the material

➢ typical value of resistivity for a conducting metal: $\rho \sim 10^{-8}$ Ω m

Example 12.9 A current of 0.10 A is driven through a heater when connected to a supply voltage of 220 V. (a) Find the resistance of the heater. (b) If the heater is made from a wire of resistivity 1.8×10^{-5} Ω m and a diameter of 0.24 mm, find the length of the wire.

✍
$$R = \frac{V}{I} = \frac{220}{0.10} \quad \Rightarrow \quad R = 2200\ \Omega$$

$$R = \frac{\rho L}{A} \quad \Rightarrow \quad L = \frac{RA}{\rho} = \frac{2200 \times \pi \times (0.12 \times 10^{-3})^2}{1.8 \times 10^{-5}} \quad \Rightarrow \quad L \approx 5.5 \text{ m} \qquad \square$$

Example 12.10 A uniform wire of resistance R has a length of L and a cross-sectional area of A. If the wire is stretched to twice the length while its volume remains constant, what is the new resistance R' in terms of R?

✍ volume: $V = LA =$ constant, so $A \propto \frac{1}{L}$, hence doubling L means A would become halved

resistance: $R = \frac{\rho L}{A} \propto \frac{L}{A} \quad \Rightarrow \quad \frac{R'}{R} = \frac{L'}{L} \times \frac{A}{A'} = 2 \times 2 \quad \Rightarrow \quad R' = 4R \qquad \square$

12.1.5 electrical power

current flowing through any electrical component transforms electrical energy into other forms

rate at which electrical energy is transferred, or electrical power, is computed as

$$P = \frac{\Delta W}{\Delta t} \xlongequal{V=\frac{W}{Q}} \frac{\Delta Q \cdot V}{\Delta t} \xRightarrow{I=\frac{\Delta Q}{\Delta t}} \boxed{P = IV}$$

substituting $V = IR$, formula for electrical power also takes two useful forms

$$\boxed{P = I^2 R} \quad \text{or} \quad \boxed{P = \frac{V^2}{R}}$$

➢ unit for electrical power: $[P] = \mathrm{W}$ (watt)

if a p.d. of 1 V drives a current of 1 A, power produced is 1 W, i.e., $1\ \mathrm{W} = 1\ \mathrm{A} \cdot 1\ \mathrm{V}$

Example 12.11 An electric toaster is labelled '220 V 900 W'. When it is operating normally, find (a) the current through it, (b) the resistance in the toaster.

✍ current: $I = \frac{P}{V} = \frac{900}{220} \quad \Rightarrow \quad I \approx 4.09\ \mathrm{A}$

resistance: $R = \frac{V}{I} = \frac{220}{4.09}$, or alternatively, $R = \frac{V^2}{P} = \frac{220^2}{900} \quad \Rightarrow \quad R \approx 53.8\ \Omega$ □

Example 12.12 Two conducting cylinders X and Y are of the same diameter and the same material. The length of the cylinder X is twice of that of the cylinder Y. If the two cylinders are connected in series to a power supply so that the current through them are equal, find the ratio of the electrical powers dissipated in the two cylinders $\frac{P_X}{P_Y}$.

✍ for either cylinder, electrical power: $P = I^2 R = I^2 \cdot \frac{\rho L}{A} = I^2 \cdot \frac{4\rho L}{\pi d^2}$

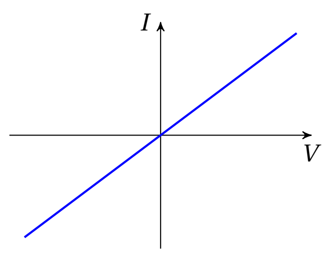
same current I, same diameter d, same resistivity ρ, so power: $P \propto L \quad \Rightarrow \quad \frac{P_X}{P_Y} = \frac{L_X}{L_Y} = 2$ □

12.2 electrical components & I-V characteristics

different components behave differently when a p.d. is applied

the behaviours can be shown in an I-V characteristics graph

12.2.1 ohmic conductors



resistance of an ohmic conductor is constant for all currents

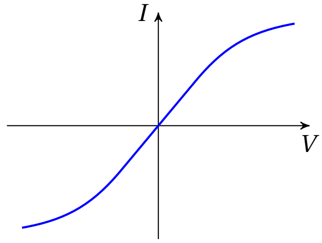
so current is directly proportional to p.d.

➢ metals wires are usually considered as ohmic conductors

resistance of metal wires is nearly constant at fixed temperature

➢ resistance of ohmic conductor equals inverse of the gradient

12.2.2 lamp filaments



lamp filaments are usually made from tungsten

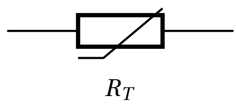
for small currents, filament behaves like an ohmic component

for larger currents, filaments are heated to high temperature[90]

resistance of filament would increase as V or I increases

➢ to find resistance of lamp filament, we read off values of I and V, then the ratio $\frac{V}{I}$ (instead of inverse of gradient) gives resistance

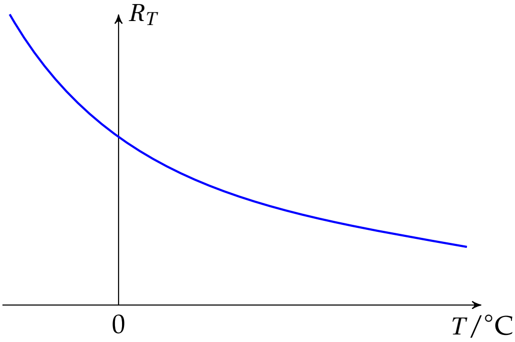
12.2.3 thermistors

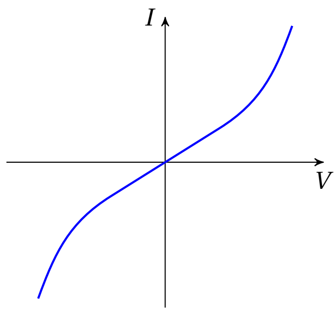


thermistor has a resistance that varies with temperature

as temperature rises, resistor of thermistor becomes lower[91][92]

at greater currents, current increases more for same p.d. increase





➢ to find resistance of thermistor, we also compute the ratio $\frac{V}{I}$

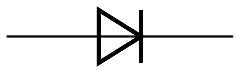
[90]This is because the vibration of the metal lattice would increase as the temperature goes up, free electrons are more likely to be scattered, so the metal becomes less conductive, the current does not increase as much for the same p.d. increase.

[91]In A-Levels, we consider *NTC (negative temperature coefficient)* thermistors only. This means their resistance decreases as temperature rises. There also exist *PTC (positive temperature coefficient)* thermistors, but let's pretend they don't exist in this course.

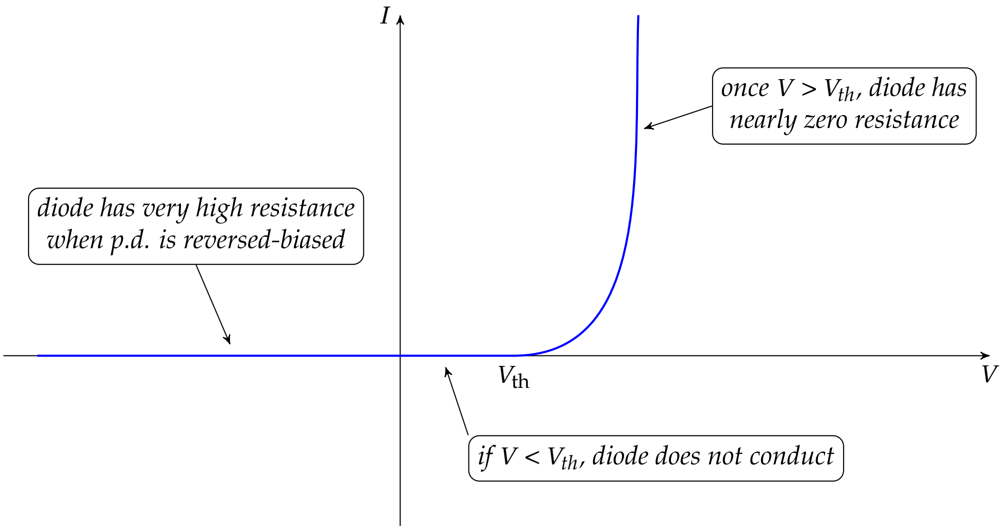
[92]The behaviour of thermistors can be explained in terms of band theory of solids, about which you are going to learn at A2 level. To put it in simple words, we can say that electrons are bound to atoms and cannot move freely at low temperatures, but they can gain energy and break free from atoms if temperature goes up. There are more free electrons available to conduct electricity, so resistance decreases.

12.2.4 diodes

diode only allows current to pass in one direction



- when an *ideal* diode is forward-biased, $R_D \to 0$
- when an *ideal* diode is reversed-biased, $R_D \to \infty$



- a practical semiconductor diode become conductive once p.d. reaches a set value V_{th}
  this value is called the diode's *threshold voltage* (about 0.6 ~ 0.7 V for a silicon diode)
  when p.d. applied is forward-biased and greater than V_{th}, diode has very low resistance
  so there is a sharp increase in current once $V > V_{\text{th}}$
- a practical diode has very high resistance if reversed-biased
  but if reverse voltage reaches the *breakdown voltage*, diode would become conductive[93]

[93]Typical breakdown voltage for a diode range from a few volts to several hundred volts, depending on its desired application.

Chapter 13

Current Electricity

13.1 Kirchhoff's circuit laws

Kirchhoff's laws are two rules that deal with the current and voltage in a circuit

they were first formulated by German physicist *Gustav Robert Kirchhoff* in 1845

the two rules allow for analysis of complex circuits in the field of electrical engineering

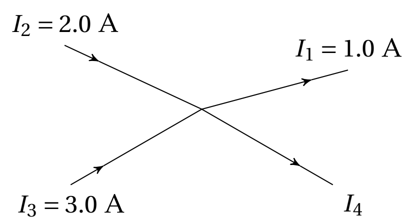
13.1.1 first law

Kirchhoff's first law, also called **Kirchhoff's current law** (KCL), states that for any point in a circuit, sum of currents flowing in equals sum of currents going out: $\boxed{\sum I_{\text{in}} = \sum I_{\text{out}}}$

➢ recall currents are produced by movement of electric charges

so Kirchhoff's first law is therefore a consequence of *conservation of electric charges*

Example 13.1 The diagram shows part of an electric circuit. The values of I_1, I_2 and I_3 are labelled. Calculate the current I_4.



✍ applying the first law: $I_2 + I_3 = I_1 + I_4$

$\Rightarrow \quad I_4 = 2.0 + 3.0 - 1.0 = 4.0$ A □

13.1.2 second law

Kirchhoff's second law, also known as **Kirchhoff's voltage law** (KVL), states that for any closed loop in a circuit, sum of all e.m.f.'s of supplies is equal to sum of potential differences across resistors: $\boxed{\sum \mathscr{E} = \sum V}$

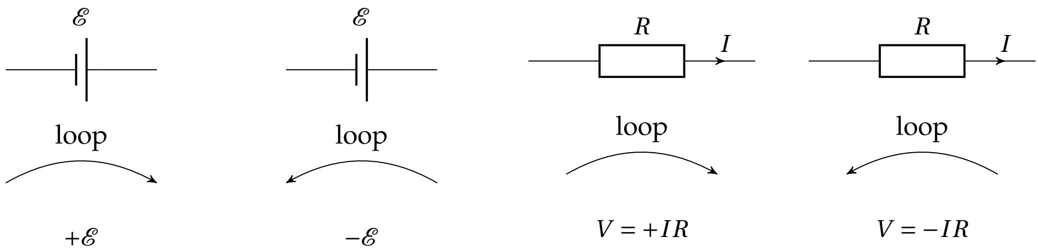
➢ recall e.m.f. and p.d. are both defined as energy transfer per unit charge e.m.f. of a supply is electrical energy produced per unit charge

p.d. across a resistor gives electrical energy consumed per unit charge

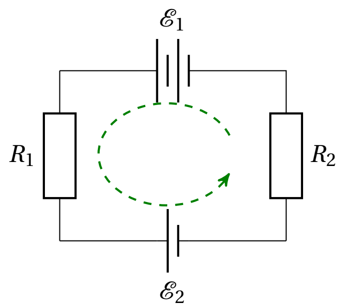
so Kirchhoff's second law is a consequence of *energy conservation*

➢ signs for $\mathscr{E}$ and V depend on the choice of loop orientation

sign conventions for $\mathscr{E}$ and V are illustrated below



Example 13.2 For the circuit shown, find an expression, in terms of $\mathscr{E}$ and R, for the electric current flowing through the resistors.

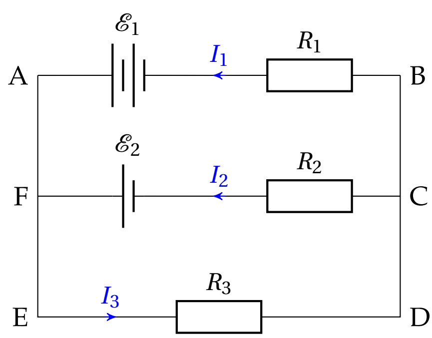


✍ apply KVL for the closed loop shown:

$$\mathscr{E}_1 - \mathscr{E}_2 = IR_1 + IR_2$$

$$I = \frac{\mathscr{E}_1 - \mathscr{E}_2}{R_1 + R_2}$$ □

Example 13.3 The diagram shows a circuit where e.m.f. of the batteries and the values of resistance are all known. Write down the equations for the currents I_1, I_2 and I_3 using Kirchhoff's laws.



✍ apply KCL for point C or F: $I_1 + I_2 = I_3$ ①

apply KVL for loop $AEDBA$: $\mathscr{E}_1 = I_1R_1 + I_3R_3$ ②

apply KVL for loop $FEDCF$: $\mathscr{E}_2 = I_2R_2 + I_3R_3$ ③

in principle, I_1, I_2 and I_3 can be solved from the three equations [94]

one can also apply KVL for loop $AFCBA$ to write: $\mathscr{E}_1 - \mathscr{E}_2 = I_1R_1 - I_2R_2$

this equation looks complex but is simply ②−③, which is not an independent equation

when writing down equations using KVL, choosing simpler loops gives easier equations □

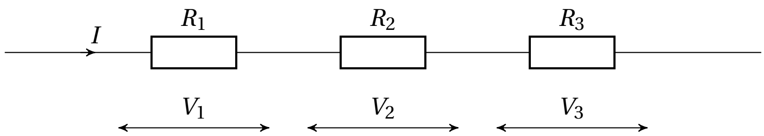
[94]The solutions are: $I_1 = \frac{(R_2+R_3)\mathscr{E}_1 - R_3\mathscr{E}_2}{(R_1+R_3)(R_2+R_3) - R_3^2}$, $I_2 = \frac{(R_2+R_3)\mathscr{E}_2 - R_3\mathscr{E}_1}{(R_1+R_3)(R_2+R_3) - R_3^2}$, $I_3 = \frac{R_1\mathscr{E}_2 + R_2\mathscr{E}_1}{(R_1+R_3)(R_2+R_3) - R_3^2}$

13.2 resistor networks

for a network of resistors, it is useful to treat the combination as a single resistor

we can derive formula for series and parallel resistors using Kirchhoff's laws

13.2.1 series resistors



three resistors in series

take three series resistors as shown

current through each resistor is the same: $I = I_1 = I_2 = I_3$

p.d. is shared between the three resistors: $V_{\text{total}} = V_1 + V_2 + V_3$

divide both sides by I, we have: $\dfrac{V_{\text{total}}}{I} = \dfrac{V_1}{I} + \dfrac{V_2}{I} + \dfrac{V_3}{I}$

so combined resistance for the three resistors is: $R_{\text{total}} = R_1 + R_2 + R_3$

in general, if there are n resistors in series, then: $\boxed{R_{\text{total}} = R_1 + R_2 + \cdots + R_n}$

- ➢ if one of the resistors in a series network increases, then R_{total} would increase
- ➢ adding an additional resistor to a network in series, then R_{total} would increase
- ➢ if n identical resistors R_0 are connected in series, then $R_{\text{total}} = nR_0$

13.2.2 parallel resistors

let's next take three parallel resistors

same p.d. across each resistor: $V = V_1 = V_2 = V_3$

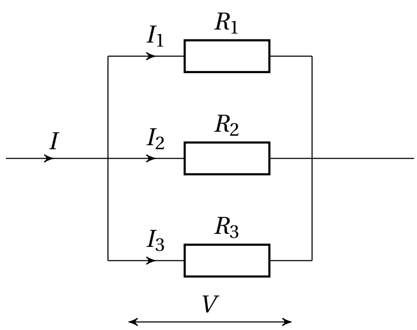
current is shared: $I_{\text{total}} = I_1 + I_2 + I_3$

divide both sides by V: $\dfrac{I_{\text{total}}}{V} = \dfrac{I_1}{V} + \dfrac{I_2}{V} + \dfrac{I_3}{V}$

so combined resistance has: $\dfrac{1}{R_{\text{total}}} = \dfrac{1}{R_1} + \dfrac{1}{R_2} + \dfrac{1}{R_3}$

in general, for n resistors in parallel:

$$\boxed{\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n}}$$



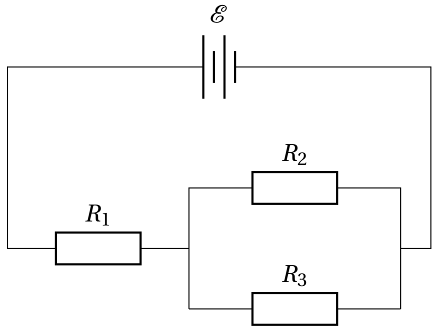
three resistors in parallel

- ➢ if one of the resistors in a parallel network increases, then R_{total} would increase

➢ adding an additional resistor to a network in parallel, then R_{total} would decrease

➢ if n identical resistors R_0 are connected in parallel, then $R_{\text{total}} = \frac{1}{n}R_0$

Example 13.4 In the circuit shown, the battery has an e.m.f. $\mathscr{E} = 18$ V, and the resistors have $R_1 = 6.0$ Ω, $R_2 = 4.0$ Ω and $R_3 = 12$ Ω. (a) Find the total resistance of all external resistors. (b) Find the current through the battery. (c) Find the p.d. across R_1, R_2 and R_3. (d) Find the total power dissipated in R_2 and R_3.



✍ total resistance: $R = R_1 + \left(\frac{1}{R_2} + \frac{1}{R_2}\right)^{-1} = 6.0 + \left(\frac{1}{4.0} + \frac{1}{12}\right)^{-1} = 6.0 + 3.0 = 9.0\ \Omega$

current through battery: $I = \frac{\mathscr{E}}{R} = \frac{18}{9.0} = 2.0$ A

p.d across R_1: $V_1 = IR_1 = 2.0 \times 6.0 = 12$ V

p.d across R_2 and R_3: $V_2 = V_3 = \mathscr{E} - V_1 = 18 - 12 = 6.0$ V

power dissipated in R_2: $P_2 = \frac{V_2^2}{R_2} = \frac{6.0^2}{4.0} = 9.0$ W

power dissipated in R_3: $P_3 = \frac{V_3^2}{R_3} = \frac{6.0^2}{12} = 3.0$ W □

13.3 practical circuits

13.3.1 power supplies & internal resistance

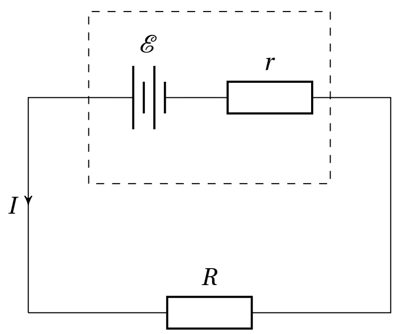
all real power sources have *internal resistance*

examples are resistance in electrolytic solution of a battery or in coils for a generator

a practical battery can be thought as the combination of an ideal battery with e.m.f. $\mathscr{E}$ and internal resistance r

if this battery is connected to a circuit of external resistance R, applying the Kirchhoff's laws, we write:

$$\boxed{\mathscr{E} = V_R + V_r = I(R + r)}$$



V_R is called the **terminal p.d.** across the battery

V_r is called the **lost volts** in the internal resistance

➢ **internal resistance** of a battery can be defined as ratio of lost volts to current in battery

➢ when a current flows through a battery, terminal p.d. will be less than battery's e.m.f.

this is because some voltage is lost due to internal resistance

➢ current in the circuit is given by: $I = \dfrac{\mathscr{E}}{R+r}$

– maximum current when terminals are shorted-out, i.e., when external load $R = 0$

greatest current is therefore: $I_{\text{max}} = \dfrac{\mathscr{E}}{r}$

though this current is limited to some finite value, this still causes great heating effects

this should be avoided because battery may be destroyed if temperature gets too high

– battery drives no current for an open circuit, i.e., $I \to 0$ if external resistance $R \to \infty$

in this case, there is no lost volts, so terminal p.d. equals e.m.f.

➢ note that there are several different notions of electrical powers

– total power generated in battery is: $P_{\text{total}} = I\mathscr{E} = I^2(R+r)$

– power delivered to external circuit is: $P_R = IV_r = I^2R$

– power dissipated (rate of thermal energy produced) in battery is: $P_r = IV_r = I^2r$

Example 13.5 A car battery has an e.m.f. of 20 V and an internal resistance of 0.50 Ω. When a starter motor of resistance of 3.50 Ω is connected to the battery, find (a) the current supplied to the motor, (b) the p.d. across the battery terminals, (c) the power at which the motor operates.

✍ current in circuit: $I = \dfrac{\mathscr{E}}{R+r} = \dfrac{20}{3.50+0.50} = 5.0$ A

terminal p.d.: $V_R = IR = 5.0 \times 3.50 = 17.5$ V

power of motor: $P_R = IV_R = 5.0 \times 17.5 = 87.5$ W or $P_R = I^2R = 5.0^2 \times 3.50 = 87.5$ W □

Example 13.6 A battery of e.m.f. 12 V is connected to a network of total resistance of 22 Ω. The current through the battery is 0.50 A. Find (a) the internal resistance of the battery, (b) the total power produced by battery, (c) the power dissipated in battery.

✍ lost volts: $V_r = \mathscr{E} - IR = 12 - 0.50 \times 22 = 1.0$ V

internal resistance: $r = \dfrac{V_r}{I} = \dfrac{1.0}{0.50} = 2.0\ \Omega$

power produced by battery: $P_{\text{total}} = I\mathscr{E} = 0.50 \times 12 = 6.0$ W

power dissipated in battery: $P_r = IV_r = 0.50 \times 1.0 = 0.50$ W or $P_r = I^2r = 0.50^2 \times 2.0 = 0.50$ W □

Example 13.7 An electric heater is operating at a working p.d. of 200 V. A current of 5.0 A is driven by a voltage source of 230 V. The source has an internal resistance of 4.0 Ω and the resistance of the connecting wires is not negligible. Find (a) the lost volts in the source, (b) the resistance of the wires, (c) the useful power from the source, (d) the efficiency of this circuit.

✍ lost volts: $V_r = Ir = 5.0 \times 4.0 = 20$ V

p.d. across wires: $V_{\text{wire}} = \mathscr{E} - V_{\text{heater}} - V_r = 230 - 200 - 20 = 10$ V resistance of wires: $R_{\text{wire}} = \dfrac{V_{\text{wire}}}{I} = \dfrac{10}{5.0} = 2.0\ \Omega$

useful power: $P_{\text{useful}} = P_{\text{heater}} = IV_{\text{heater}} = 5.0 \times 200 = 1000$ W

efficiency: $\eta = \dfrac{P_{\text{useful}}}{P_{\text{total}}} = \dfrac{IV_{\text{heater}}}{I\mathscr{E}} = \dfrac{200}{230} \approx 87\%$ □

Example 13.8 A cell of e.m.f. $\mathscr{E}$ and internal resistance r is connected in series with a variable resistor R. R is gradually increased from zero. (a) Suggest how the p.d. across the battery terminals change when R is increased. (b) For larger values of R, power delivered to R decreases. Suggest the advantage, despite the low power output, of using this cell in a circuit of larger resistance.

✍ as R increases, current in circuit decreases

less voltage is lost in cell, so terminal p.d. will increase

for same reason, less power is lost in cell, so higher efficiency for the circuit

more explicitly, terminal p.d. $V_R = IR = \dfrac{\mathscr{E}R}{R+r}$, and efficiency $\eta = \dfrac{P_R}{P_{\text{total}}} = \dfrac{I^2R}{I^2(R+r)} = \dfrac{R}{R+r}$

from these we can tell: if $R \nearrow$, then $\dfrac{R}{R+r} \nearrow$, so $V_R \nearrow$ and $\eta \nearrow$ □

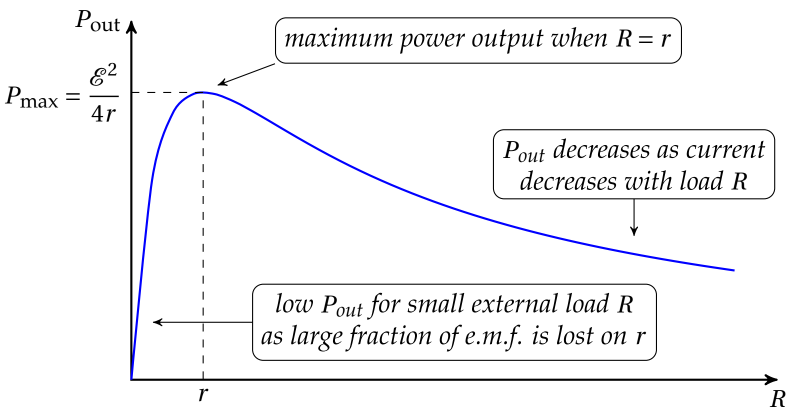
power output from a practical battery

power output to external components is

$$P_{\text{out}} = I^2R = \left(\frac{\mathscr{E}}{R+r}\right)^2 R$$

assuming $\mathscr{E}$ and r are constant, then P_{out} only depends on the external load R of the circuit

the diagram below show how P_{out} varies with R



variation of output power P_{out} from a power supply to an external resistor R

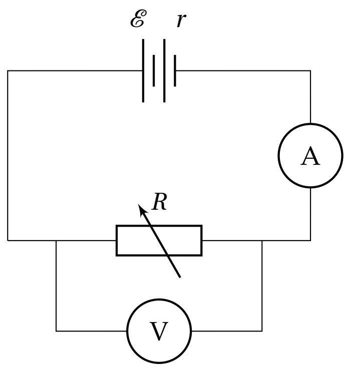
➢ when R is small, only a small fraction of e.m.f. is output as terminal p.d.

so most of the power generated is lost on the internal resistance of the source

- as R gradually increases, terminal p.d. increases, causing a temporary rise in P_{out}
  however, increase in terminal p.d. is compensated by a reduction in electric current
- when R continues to increase, current becomes so small that P_{out} gradually tends to zero
- maximum power output is archived when $R = r$, this power is given by: $P_{\text{out,max}} = \dfrac{\mathscr{E}^2}{4r}$ [95]

measurement of e.m.f. and internal resistance of a power supply

the circuit for determining e.m.f. and internal resistance of an unknown power source is shown on the right



terminal p.d. V_R is measured by the voltmeter

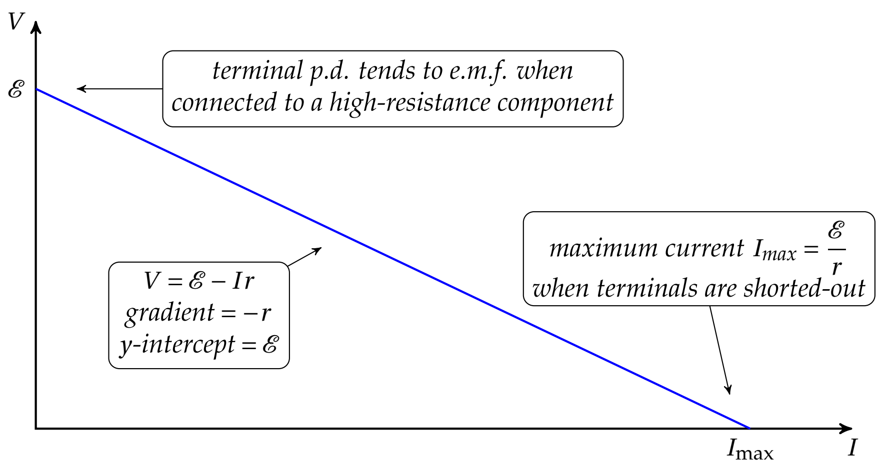
current I in circuit is measured by the ammeter

one can vary R to obtain a set of measurements for I and V

values of V can then be plotted against values of I

since $V_R = \mathscr{E} - Ir$, data points should fall in a straight line

- $\mathscr{E}$ is y-intercept of the graph
- r is given by the negative gradient



variation of terminal p.d across a battery against the current flowing through it

[95] For any two positive numbers x and y, $x + y \geq 2\sqrt{xy}$ where the equality holds if and only if $x = y$. We then have $P_{\text{out}} = \dfrac{\mathscr{E}^2 R}{(R+r)^2} = \dfrac{\mathscr{E}^2}{R + \frac{r^2}{R} + 2r} \leq \dfrac{\mathscr{E}^2}{2\sqrt{R \cdot \frac{r^2}{R}} + 2r} = \dfrac{\mathscr{E}^2}{4r}$. To obtain the greatest power, the condition $R = \dfrac{r^2}{R}$, i.e., $R = r$ must be satisfied.

13.3.2 practical ammeters & voltmeters

when connected into a circuit, ideal ammeters/voltmeters do not affect original currents

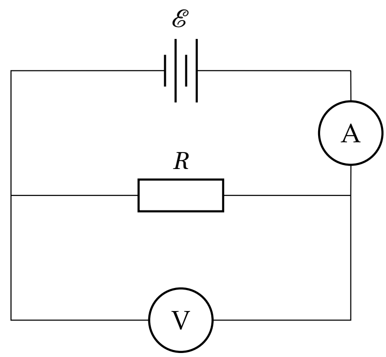
so ideal ammeter has zero resistance, and ideal voltmeter has infinite resistance

but in practice, ammeters have non-zero resistance, voltmeter have finite resistance

to find current through an ammeter/p.d. a voltmeter, treat them as normal resistor

reading on ammeter/voltmeter gives the current through/p.d. across itself

Example 13.9 The diagram shows a simple circuit. The battery has a negligible internal resistance and an e.m.f. of 8.0 V. The resistor has a fixed resistance of 20 Ω. (a) Find the readings shown on the ammeter and the voltmeter if both meters are ideal. (b) Instead, the ammeter has a non-zero resistance of 1.0 Ω and the voltmeter has a resistance of 60 Ω, what are the true readings displayed on the two meters?



✍ if both meters are ideal, then $I = \dfrac{\mathscr{E}}{R} = \dfrac{8.0}{20} = 0.40$ A, and $V = \mathscr{E} = 8.0$ V

for non-ideal case, total resistance in circuit: $R_{\text{total}} = R_A + \left(\dfrac{1}{R} + \dfrac{1}{R_V}\right)^{-1} = 1.0 + \left(\dfrac{1}{20} + \dfrac{1}{60}\right)^{-1} = 16\ \Omega$

current through ammeter: $I = \dfrac{\mathscr{E}}{R_{\text{total}}} = \dfrac{8.0}{16} = 0.50$ A

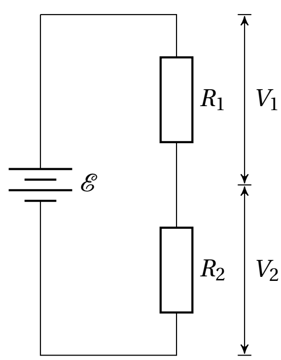
p.d. across voltmeter: $V = \mathscr{E} - V_A = \mathscr{E} - IR_A = 8.0 - 0.50 \times 1.0 = 7.5$ V □

13.3.3 potential dividers

one type of useful circuit is the potential divider

a **potential divider** can produce an output voltage that is a fraction of input voltage

a typical potential divider circuit is shown



p.d. across R_1 and R_2 satisfy the relationship:

$$\frac{V_1}{V_2} = \frac{IR_1}{IR_2} \quad \Rightarrow \quad \frac{V_1}{V_2} = \frac{R_1}{R_2}$$

we also have the relation: $V_1 + V_2 = \mathscr{E}$

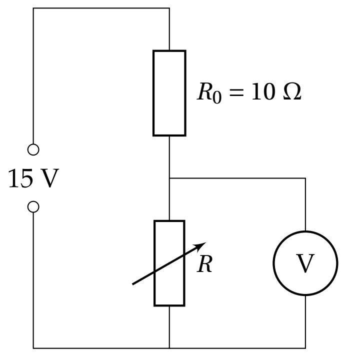
these together give the *potential divider equation*:

$$\boxed{V_1 = \frac{R_1}{R_1 + R_2} \times \mathscr{E}} \qquad \boxed{V_2 = \frac{R_2}{R_1 + R_2} \times \mathscr{E}}$$

this means R_1 and R_2 divide up the p.d. supplied to them

proportion of p.d. share depends on relative resistance values

Example 13.10 The diagram shows an electric circuit incorporating a power supply of negligible internal resistance and a high-resistance voltmeter. Determine the range of voltage reading on the voltmeter as the variable resistor R is adjusted over its full range from 0 Ω to 50 Ω.



✍ $V_{\text{min}} = \dfrac{R_{\text{min}}}{R_{\text{min}} + R_0} \times V_{\text{in}} = \dfrac{0}{0+10} \times 15 \quad \Rightarrow \quad V_{\text{min}} = 0\ \text{V}$

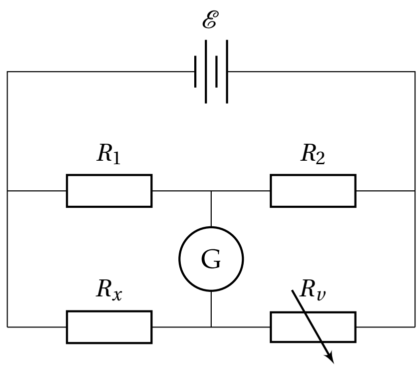
$V_{\text{max}} = \dfrac{R_{\text{max}}}{R_{\text{max}} + R_0} \times V_{\text{in}} = \dfrac{50}{50+10} \times 15 \quad \Rightarrow \quad V_{\text{max}} = 12.5\ \text{V}$

so voltage reading ranges from 0 to 12.5 V □

bridge circuits

bridge circuits can be designed by altering the potential divider circuit [96]

by balancing two legs of a bridge circuit, an unknown resistance can be measured



the circuit diagram shows a typical bridge circuit

R_x is the unknown resistance to be measured

resistance of R_1, R_2 and variable resistor R are known

R is adjusted until no current flows through galvanometer, the bridge is then said to be *balanced*

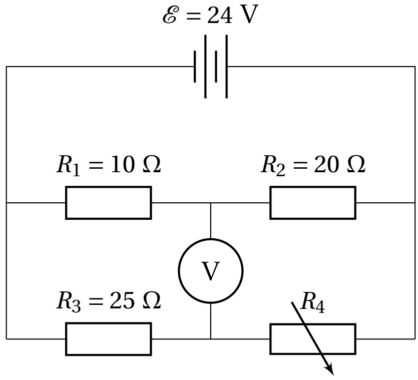
p.d. share between R_1 and R_2 is same as p.d. share between R_x and R: $\dfrac{R_1}{R_2} = \dfrac{R_x}{R}$

so resistance of R_x can be determined to great accuracy [97]

[96]The bridge circuit we will be looking at is known as the *Wheatstone bridge*. The circuit design was invented by British scientist *Samuel Hunter Christie* in 1833 and later improved by another British scientist *Charles Wheatstone* in 1843.

[97]Measurement with the Wheatstone bridge circuit can be extremely accurate because the scheme illustrates the concept of a difference measurement, which can be done to very high accuracy.

Example 13.11 Four resistors are connected to a battery as shown. (a) If the variable resistor R_4 is adjusted to have a resistance of 35 Ω, what is the reading on the voltmeter? (b) If the voltmeter reads zero, what is the resistance for R_4?



✍ (a) p.d. across R_1: $V_1 = \dfrac{10}{20+10} \times 24 = 8.0$ V

p.d. across R_3: $V_3 = \dfrac{25}{35+25} \times 24 = 10.0$ V

voltmeter reading: $V = V_3 - V_1 = 10.0 - 8.0 = 2.0$ V

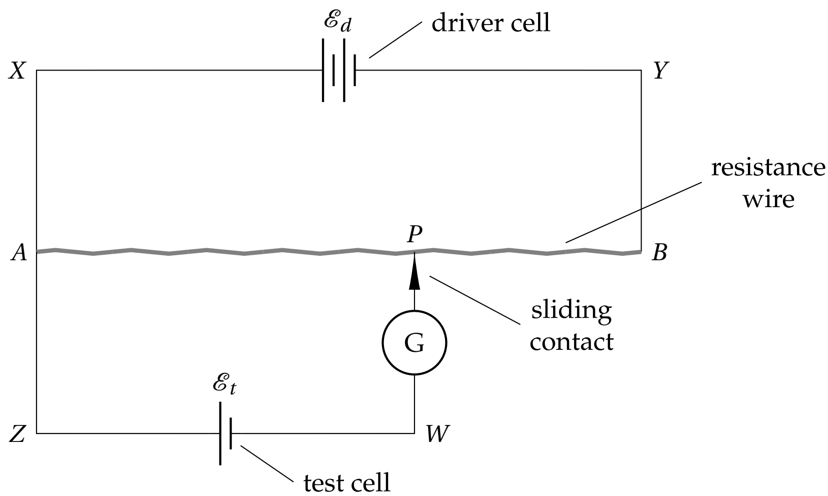
(b) $V = 0$ means p.d. share between R_1 and R_2 is same as p.d. share between R_3 and R_4

$$\frac{R_1}{R_2} = \frac{R_3}{R_4} \quad \Rightarrow \quad \frac{10}{20} = \frac{25}{R_4} \quad \Rightarrow \quad R_4 = 50\ \Omega$$ □

potentiometer

another useful type of potential divider is the *potentiometer*

p.d.'s or e.m.f.'s can be compared in terms of length quantities with a potentiometer



the diagram shows a potentiometer being used to measure the e.m.f. of a unknown cell

suppose driver cell has no internal resistance and an e.m.f. of $\mathcal{E}_d$ that is already known

to find e.m.f. of test cell, we adjust position of contact P until galvanometer shows zero

taking loop $XAPBYX$, we have: $\mathcal{E}_d = V_{AB} = IR_{AB}$

taking loop $ZAPWZ$, we have: $\mathcal{E}_t = V_{r,t} + V_{AP} = \cancelto{0}{ir_t} + IR_{AP} = IR_{AP}$

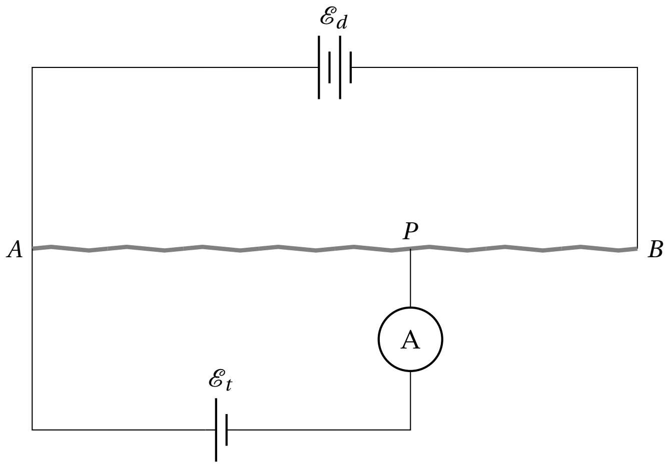
compare the two equations, we find: $\dfrac{\mathscr{E}_t}{\mathscr{E}_d} = \dfrac{R_{AP}}{R_{AB}}$

recall resistance of uniform wire is proportional to its length: $\dfrac{R_{AP}}{R_{AB}} = \dfrac{AP}{AB}$

now ratio of e.m.f.'s of the two cells are related to ratio of two lengths AP and AB

e.m.f. of test cell is given by: $\mathscr{E}_t = \dfrac{AP}{AB} \times \mathscr{E}_d$

Example 13.12 In the circuit below, the driver cell has an e.m.f. of 17 V and an internal resistance of 5.0 Ω. AB is a resistance wire of total resistance 80 Ω and length 80 cm. The e.m.f. of a test cell with an unknown internal resistance is to be determined. The moving contact P is adjusted to a position where $AP = 50$ cm such that the ammeter shows no reading. (a) Find the current flowing in the resistance wire. (b) Find the p.d across A and P. (c) State the e.m.f. of the test cell.



✍ current in wire: $I = \dfrac{\mathscr{E}}{R_{AB} + r_d} = \dfrac{17}{80 + 5.0} = 0.20$ A

resistance of AP: $R_{AP} = \dfrac{AP}{AB} \times R_{AB} = \dfrac{50}{80} \times 80 = 50\ \Omega$

p.d. across AP: $V_{\mathrm{AP}} = IR_{\mathrm{AP}} = 0.20 \times 50 = 10$ V

note that there is no need to worry about internal resistance of the test cell

since no current flows through test cell at point of balance, so no lost volts

hence, e.m.f. of test cell: $\mathscr{E}_t = \cancelto{0}{V_{r,t}} + V_{AP} = 10$ V □

Chapter 14

Electric Fields

14.1 electric forces & fields

14.1.1 electrostatic forces

any charged object can interact with any other charged object

this force is known as the *electrostatic force*, or in short, the electric force

➢ electric force can be *repulsive* or *attractive*, depending on nature of charge

like charges would repel, opposite charges would attract

14.1.2 electric fields

charged objects can interact without physically touching each other

to explain this *action at a distance*, we introduce the concept of *force fields*

a force field is basically a region of space where a body experiences a force [98]

electric field is a region of space in which a charged object is acted by an electric force

to elaborate, any charged body can create an electric field around it

property of space is altered, so another charged body in the field experiences a force

from this point of view, when charged bodies interact, they do not directly exert forces upon each other, but they interact through electric fields

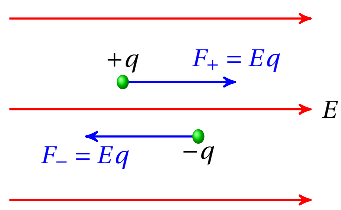
[98] Apart from electric fields (act on bodies carrying electric charges), other examples of force fields in physics are gravitational fields (act on objects with mass) and magnetic fields (act on magnets, electric currents or moving charges). You will be studying these subjects at A-Levels.

14.1.3 electric field strength

to describe how strong an electric field is at a point, we define the *electric field strength*

> **electric field strength** is the electric force acting per unit positive charge: $\boxed{E = \frac{F}{q}}$

- SI unit of electric field strength: $[E] = \mathrm{N\,C^{-1}} = \mathrm{V\,m^{-1}}$
- electric field strength is a vector quantity
  direction of field strength is defined as the force experienced by a *positive* test charge
- given the field strength E, electric force on a test charge q can be determined
  magnitude of electric force is given by $F = Eq$
  direction of electric force depends on both direction of E and polarity of q
  – for positive charge, F is in same direction as E
  – for negative charge, F is in opposite direction as E



- electric fields due to two (or more) different sources can overlap
  the combined field strength is *vector sum* of the individual fields

14.1.4 electric field lines

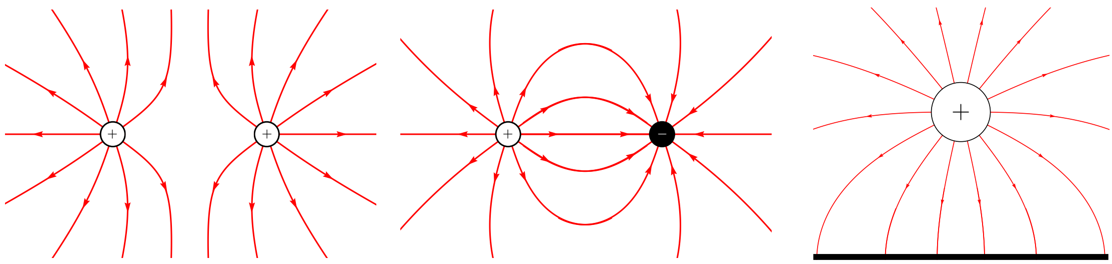
pattern of an electric field can be graphically represented by **electric field lines**

- field lines can give information about electric field strength
  – tangents of the field lines point in same direction as field strength
  – density of the lines gives information about magnitude of field strength
- some rules for drawing electric field lines are given below
  – field line starts radially outwards from a positive charge (or from infinity)
  – field line ends up radially inwards on a negative charge (or to infinity)
  – field lines can never intersect each other[99]
  – filed lines are always perpendicular to the surface of a conductor[100]

[99]This is because E at a given position must point in a definite direction unless it is zero.
[100]This is because $E = 0$ everywhere inside a conductor, so field strength cannot have components that are tangent to a conducting surface.

Example 14.1 The diagrams below show the pattern of various electric fields.

radial fields around (a) an isolated positive charge, (b) an isolated negative charge



field between (c) two equal and like charges, (d) two equal but opposite charges,

(e) a positive charge and a large conducting sheet that is earthed

14.2 uniform electric fields

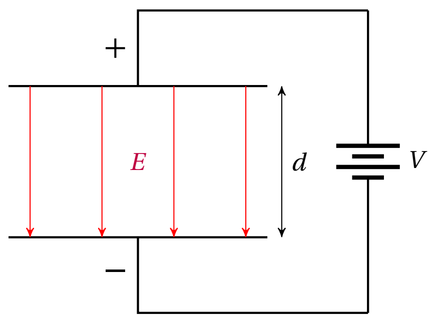
let's take two parallel metal plates and connect them to a high-voltage supply as shown

a charge in the region between the plates experiences a constant force regardless of its location in the field

field strength is constant throughout the region

this field is said to be a **uniform electric field**

field lines of a uniform field are a set of parallel lines of equal spacing as shown on the right



➢ if the two plates are separated by a distance of d, and a p.d. of V is applied across them

then magnitude of field strength of the uniform field is given by: $\boxed{E = \frac{V}{d}}$

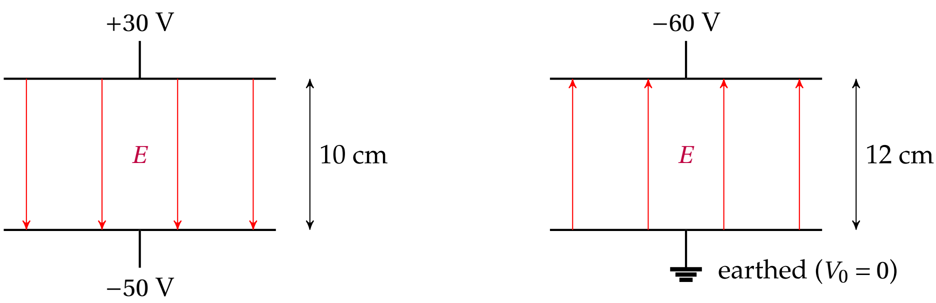
proof: work done to bring a test charge q from positive plate to negative plate: $W = Fd = Eqd$

the change in electric potential energy is: $\Delta E_p = qV$ (check §12.1.3 for meaning of p.d.)

work-energy theorem tells us these two must be equal: $Eqd = qV$, so we find: $E = \dfrac{V}{d}$

➢ electric field between metal plates points from high potential towards low potential

Example 14.2 The following examples show the set-up of two uniform electric fields.



$$E = \frac{V}{d} = \frac{(+30) - (-50)}{0.10} = 800\ \mathrm{N\,C^{-1}} \qquad E = \frac{V}{d} = \frac{0 - (-60)}{0.12} = 500\ \mathrm{N\,C^{-1}}$$

Example 14.3 Two horizontal parallel plate conductors are separated by 1.5 cm in air. A potential difference of 36 V is applied across the plates. Compare the electric force and gravitational force acting on a proton between the plates.

✍ electric field strength: $E = \dfrac{V}{d} = \dfrac{36}{1.5 \times 10^{-2}} = 2400\ \mathrm{N\,C^{-1}}$

electric force: $F_E = Eq = 2400 \times 1.60 \times 10^{-19} \approx 3.84 \times 10^{-16}$ N

weight of the proton: $W = mg = 1.67 \times 10^{-27} \times 9.81 \approx 1.64 \times 10^{-26}$ N

this shows electric force on proton are much stronger than gravitational forces

when dealing with motion of sub-atomic particles (e.g., protons, electrons, etc.) in an electric field, effects of gravity can therefore be ignored □

14.2.1 equilibrium between electric force and weight

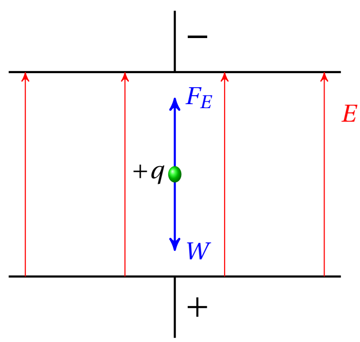
a charged particle (e.g., a dust particle or an oil droplet) can be held stationary in a uniform electric field

electric force is in equilibrium with weight of particle

$$F_E = W \quad \Rightarrow \quad Eq = mg$$

where field strength E is given by: $E = \dfrac{V}{d}$

note that equilibrium is possible only if F_E acts in opposite direction to W, so direction of the applied field depends on polarity of the charged particle



Example 14.4 A uniform electric field is set up between the two parallel oppositely-charged metal plates separated by 8.0 cm. An oil droplet of mass 0.040 g and charge −20 nC is placed in the field. The droplet stays at rest. (a) Find the strength and the direction of the electric field. (b) What is the voltage required to produce this field? (c) If the separation between the plates is reduced, what would happen to the oil droplet?

✍ (a) equilibrium so electric force equals weight: $F_E = mg = 0.050 \times 10^{-3} \times 9.81 \approx 3.92 \times 10^{-4}$ N

field strength: $E = \frac{F_E}{q} = \frac{3.92 \times 10^{-4}}{20 \times 10^{-9}} \approx 1.96 \times 10^4 \text{ N C}^{-1}$

F_E must act upwards for equilibrium, but droplet is negatively-charged

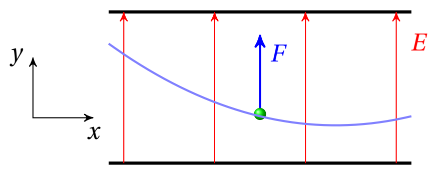
so electric field acts in the downward direction

(b) voltage between the plates: $V = Ed = 1.96 \times 10^4 \times 8.0 \times 10^{-2} \approx 1.57 \times 10^3$ V

(c) smaller separation means greater field strength, so greater electric force

$F_E > mg$ so resultant force acts upwards, droplet will accelerate upwards □

14.2.2 deflection of charged particle in uniform fields



charged particle travelling in a uniform electric field would in general undergo a *projectile-like* motion

we set up the coordinates such that electric field is along the y-axis, and x-axis is in the normal direction

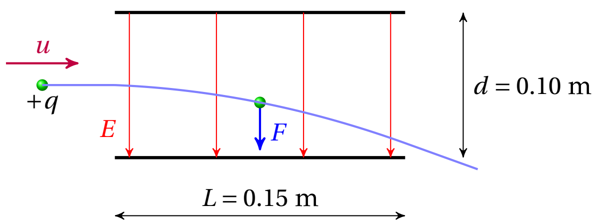
we denote x_0, y_0 and u_x, u_y as initial displacements and initial velocities at $t = 0$

electric force only acts in y-direction, no force acts in x-direction

- in x-direction, particle maintains a constant velocity: $v_x = u_x, \ x = x_0 + u_x t$
- in y-direction, particle accelerates uniformly: $v_y = u_y + at, \ y = y_0 + u_y t + \frac{1}{2}at^2$

  where acceleration in y-direction is given by $a = \frac{F}{m} = \frac{Eq}{m}$ (assuming weight is negligible)

so motion of a charged particle in a uniform electric field is analogous with projectile motion in a uniform gravitational field (see §3.5), i.e., trajectory of charged particle is *parabolic*



Example 14.5 A proton enters a uniform field of strength $1.0 \times 10^4 \text{ N C}^{-1}$ at an initial velocity of $5.0 \times 10^5 \text{ m s}^{-1}$. The initial direction of motion is at right angles to the field. The dimensions of the field are shown in the diagram. (a) What is the time taken

for the proton to pass through the field? (b) What is the deviation in the direction of the field during this time? (c) If an electron enters the field with the same initial velocity, describe how the deflection of the electron compares with that of the proton.

✍ constant horizontal velocity, so time taken: $t = \frac{L}{u} = \frac{0.15}{5.0 \times 10^5} = 3.0 \times 10^{-7}$ s

constant acceleration in direction of field: $a = \frac{F}{m} = \frac{Eq}{m} = \frac{1.0 \times 10^4 \times 1.6 \times 10^{-19}}{1.67 \times 10^{-27}} \approx 9.58 \times 10^{11}$ m s^{-2}

displacement moved in this direction: $\Delta y = \frac{1}{2}at^2 = \frac{1}{2} \times 9.58 \times 10^{11} \times (3.0 \times 10^{-7})^2 \approx 0.043$ m

electron would experience a much greater deflection in the opposite direction

- electron has opposite charge to proton, so deflect in the opposite direction
- mass of electron is much smaller, so much greater acceleration

  time spent in the field is the same, so electron has much greater deflection □

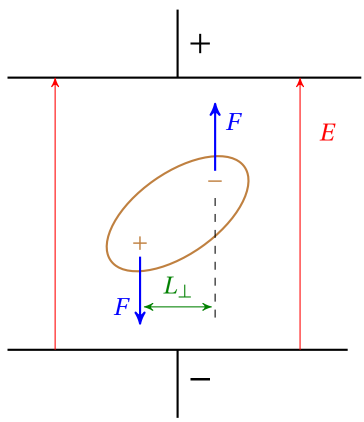
14.2.3 polar molecules in uniform electric fields

centre of positive and negative charges for a molecule do not necessarily overlap

this molecule is still electrically neutral, i.e., it has a zero net charge

such a molecule is called a *polar molecule*, and it could be affected by an electric field

the diagram shows a polar molecule in a uniform field



force on centre of positive charge and force on centre of negative charge are labelled as shown

this is a pair of equal but opposite forces

so they give rise to a resultant *torque/moment*

recall *torque of couple* is defined as one force times perpendicular distance between the force pair (see §5.1.1)

$$\tau = F_E L_\perp \quad \Rightarrow \quad \tau E q L_\perp$$

this torque would cause the polar molecule to *rotate*

CHAPTER 15

Sub-atomic Particles

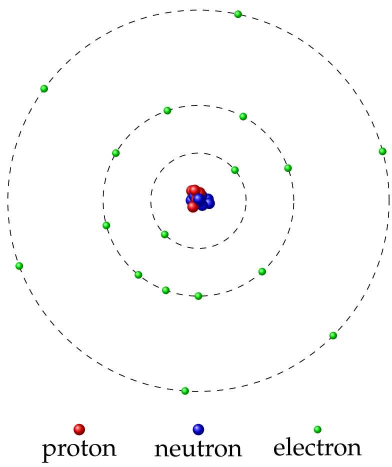
15.1 atomic structure

15.1.1 atomic model

all matter is composed of tiny particles called **atoms**

each different element has its own type of atom

the diagram below illustrates a simple atomic model



- at centre of each atom is the **nucleus**
- radius of an atom $\sim 10^{-10} \sim 10^{-9}$ m

  radius of a nucleus $\sim 10^{-15} \sim 10^{-14}$ m
- particles that make up a nucleus are called **nucleons**

  nucleons come in two types, **protons** and **neutrons**
- outside the nucleus are the **electrons**

  electrons move around the nucleus in a *cloud*
- protons, neutrons and electrons are the building blocks of all atoms and hence the building blocks for all matter
- properties of protons, neutrons and electrons are listed below
  - charge of each proton and electron is the elementary charge unit: $e = 1.60 \times 10^{-19}$ C
  - neutron has zero charge
  - a proton has very similar mass as a neutron
  - electron has much smaller mass than a nucleon (proton or neutron)

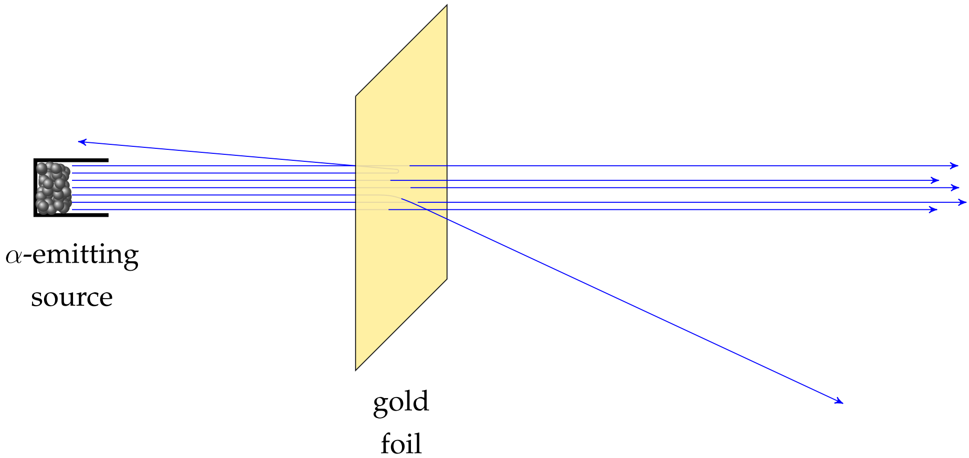
subatomic particle	charge	mass	location found
proton	$+e$	$m_p = 1.67 \times 10^{-27}$ kg	nucleus
neutron	0	$m_n = 1.67 \times 10^{-27}$ kg	nucleus
electron	$-e$	$m_e = 9.11 \times 10^{-31}$ kg	in outer atom

15.1.2 α-particle scattering experiment

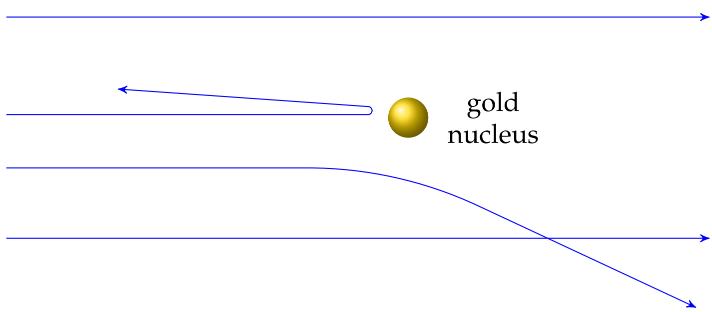
experiments were designed to verify if the model gives the right description of an atom

one of the most important experiments carried out was the **α-particle scattering experiment**

under the direction of *Ernest Rutherford*, *Hans Geiger* and *Ernest Marsden* studied atomic structure by firing α-particle beam towards a thin gold foil at the University of Manchester



Rutherford's α-particle scattering experiment



paths of α-particles as they approach and pass by a gold nucleus

- the main experimental results of the experiments are
  - most α-particles pass straight through the foil with almost no deflection
  - very few α-particles (about 1 in 10^4) are deflected by large angles.
- to explain the observations, Rutherford concluded the following about nature of atoms
  - most space in an atom is empty

(so most α-particles merely pass through empty space without deflection)

- there is a tiny positively-charged core, called the nucleus, at centre of an atom
  (so only a small fraction of α-particles would interact with the nucleus)
- almost all mass of the atom is concentrated in the nucleus
  (so interaction is influential if nucleus happens to sit on the path of α-particle beam)

based on these understandings, Rutherford proposed the atomic model introduced in §15.1.1

this model is now known as *Rutherford's model*, or *planetary model of the atom* [101] [102]

15.1.3 nuclide notation

nuclides are specific types of atoms or nuclei

a nuclide is uniquely defined by its **proton number** Z and its **nucleon number** A

a nuclide is usually denoted as $\boxed{{}^{A}_{Z}X}$, where X is the chemical symbol for the element

➢ by definition, Z gives number of protons, A gives number of nucleons

---

[101] Before Ernest Rutherford, the most popular atomic model at the time was proposed by *J. J. Thompson* of the Cavendish Lab at the University of Cambridge. Thompson first discovered electrons from cathode rays. He showed that electrons are negatively-charged and hence proved that an atom is made up of electrons and something that carries positive charges. He suggested that electrons are evenly distributed in a positive cloud of matter, which is now called the *plum-pudding model*.

To verify this model, Rutherford's group set out and designed the gold foil experiment. The plum-pudding model predicts that deflection of α-particles should be uniformly distributed around the central beam, but there is no reason that an α-particle would undergo a large change in its direction of motion, like a cannonball hitting a piece of paper and rebounding backwards. The results of the scattering experiment meant that the plum-pudding model had to be overturned. In order to explain the experimental observations, Rutherford then put forward the new model which now bears his name.

[102] Rutherford originally suggested that electrons rotate around the nucleus in *circular* orbits due to attractive electrostatic force, like the planets orbit around the sun. This is why Rutherford's model of atomic structure is also called *the planetary model of the atom*.

However, later studies showed that Rutherford's model had its own problems. Classical electromagnetic theory suggests that any charged particle in accelerated motion would radiate energy, so the orbiting electrons will gradually lose all its energy and fall into the nucleus. An atom can never be stable if the Rutherford's model is correct. The full description of the atomic structure requires the understanding of the *quantum* behaviour of electrons, which are described by *probability waves*. Electrons actually do not move in circular orbits but have a certain probability to be anywhere in space at any time, forming an *electron cloud*.

as a consequence, number of neutrons is given by $N = A - Z$

➢ proton number Z can be further interpreted as **charge number** of the particle

recall each proton carries charge $+e$, and neutrons are chargeless

so Z further determines electrical charge of the nucleus: $\boxed{Q = +Ze}$

this extension will be important when we deal with other particles later in this chapter

➢ nucleon number A can be interpreted as **mass number** of the particle

each proton and neutron is of approximately the same mass

so mass of nucleus is determined by total number of protons and neutrons: $\boxed{m = A\mathrm{u}}$

where $1\ \mathrm{u} = 1.66 \times 10^{-27}$ kg is the **unified atomic mass unit** [103]

➢ using this notation, protons, neutrons and electrons can be represented by: $\boxed{{}^1_1\mathrm{p}}$, $\boxed{{}^1_0\mathrm{n}}$, $\boxed{{}^{0}_{-1}\mathrm{e}}$

Example 15.1 The nuclide uranium-238 can be denoted by ${}^{238}_{92}\mathrm{U}$. (a) State its proton number and its neutron number. (b) Find the charge and mass of the uranium-238 nucleus.

✍ proton number: $Z = 92$

neutron number: $N = A - Z = 238 - 92 = 146$

nuclear charge: $Q = +Ze = +92 \times 1.60 \times 10^{-19} \approx 1.47 \times 10^{-17}$ C

mass of nucleus: $m = A\mathrm{u} = 238 \times 1.66 \times 10^{-27} \approx 3.95 \times 10^{-25}$ kg □

Example 15.2 Radius of a copper-64 nucleus is 4.8×10^{-15} m. Calculate the mean nuclear density.

✍
$$\rho = \frac{m}{V} = \frac{A\mathrm{u}}{\frac{4}{3}\pi r^3} = \frac{64 \times 1.66 \times 10^{-27}}{\frac{4}{3}\pi \times (4.8x10^{-15})^3} \quad \Rightarrow \quad \rho \approx 2.3 \times 10^{17}\ \mathrm{kg\ m^{-3}}$$

this is much greater than density of a copper block (about 8.9×10^3 kg m^{-3})

we can therefore infer that the nucleus indeed takes most mass of the atom □

15.1.4 chemical properties of elements

chemical properties of an atom depend on the number of electrons

➢ proton number Z also gives the **atomic number** of the element

[103]Note that the unified atomic mass (1.661×10^{-27} kg), formally defined as one twelfth of the mass of an carbon-12 atom, is slightly less than the mass of a *free* proton ($m_p = 1.673 \times 10^{-27}$ kg or that of a *free* neutron ($m_n = 1.675 \times 10^{-27}$ kg). This is because energy goes out when protons and neutrons combine to form a nucleus, and this would be associated with a decrease in the mass as suggested by *Albert Einstein's mass-energy equivalence principle*. You can think of u as the average mass of a proton or a neutron when a group of them squeeze together. If you want to find out the mass of a nucleus, you should use u. But if you are dealing with a single proton or a single neutron, you should use m_p and m_n instead.

in a neutral atom, there are same number of protons as electrons

so atoms of same chemical properties, i.e., atoms of same element have same proton number

➢ atoms of same chemical properties can have different nuclei

isotopes of same element have same number of protons but different number of neutrons

Example 15.3 State and explain whether carbon-12 ($^{12}_{6}C$) and carbon-14 ($^{14}_{6}C$) are different isotopes of the same element.

✍ both carbon-12 and carbon-14 atoms have 6 protons

but carbon-12 has 6 neutrons, carbon-14 has 8 neutrons

same proton number but different neutron number, so different isotopes of same element □

15.2 radioactivity

15.2.1 radioactive decays

if a nucleus is unstable then it may break down by releasing radiation from the nucleus

this process is called **radioactive decay**

➢ three most common types of radioactive decays are α-decay, β-decay and γ-decay

➢ radioactive decays are *spontaneous*

decay events are independent of outside conditions, such as temperature, pressure, etc.

whether one nucleus is about to decay is also independent of the other nuclei around it

radioactive decays only depend on the internal stability of the nucleus

➢ radioactive decays are *random*

one cannot predict when and which nucleus will decay

an unstable nucleus might decay at any time

whether or not decay occurs can only be given by probabilities

15.2.2 properties of α-, β- & γ-radiation

nature of α-radiation

α-particles are *helium nuclei*, each made up of two protons and two neutrons

an α-particle therefore carries a positive charge of $2e$, and a mass of 4u

α-particles emitted from a given source travel at almost the same speed

speed of α-particles is usually less than one-tenth of the speed of light

nature of β-radiation

β-radiation is a beam of high-speed *electrons*

β-particle is negatively charged and very light compared with a nucleus

β-particles travel at very high speeds close to the speed of light

nature of γ-radiation

γ-radiation, or γ-photon, is a high-frequency electromagnetic wave

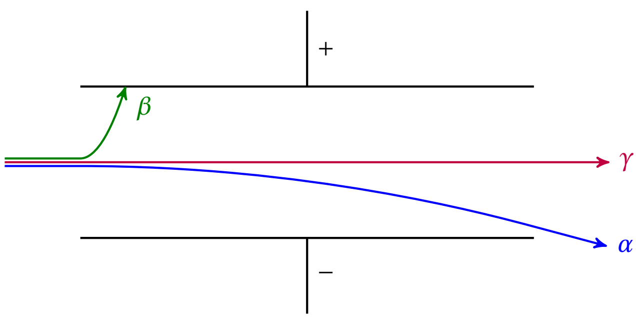
γ-radiation is electrically neutral and has no mass

like any other electromagnetic wave, γ-radiation travels at the speed of light in vacuum

brief summary for nature of α-, β- & γ-radiation

	α-radiation	β-radiation	γ-radiation
nature	helium nucleus	electron	electromagnetic wave
notation	${}^4_2\alpha$	${}^{\ 0}_{-1}\beta$	${}^0_0\gamma$
mass	4u	$\frac{1}{1800}\text{u} \approx 0$	0
charge	$+2e$	$-e$	0
speed	$0.1c$	$0.99c$	c

Example 15.4 If a single beam of α-particles, β-particles and γ-radiation is sent into a uniform electric field as shown, sketch and explain the paths of the beam.



✍ γ-radiation has zero charge, so it is not affected by electric field

α-radiation is positively-charged, so it deflects towards the negative plate

β-radiation is negatively-charged, so it deflects towards the positive plate

β-particles are much lighter, so they have greater deflection □

penetration & ionising power

different radiation have different abilities to pass through materials

α-, β- and γ- radiation are able to knock out orbital electrons from an atom

so radiation can cause the originally neutral atom to become charged

this process is called **ionising**

➢ greater ionising ability would imply lower penetration power

radiation loses energy as it passes through a substance

this energy loss is transferred to electrons, causing ionisation of atoms

so the faster radiation gives up its energy, the lower the penetration ability

➢ γ-radiation is very penetrative but has very weak ionising power

γ-radiation is not charged, and hence it does not interact strongly with electrons

a γ-photon loses all its energy in single collision and gets absorbed

i.e., one γ-photon can only ionise one atom

so it has few collisions with electrons, so γ-radiation is not very ionising

➢ α- and β-radiation more ionising than γ-radiation

this is because α- and β-radiation are charged, making them interact with electrons more easily

➢ α-particles are highly ionising

recall α-particles are much more massive than electrons

so one α-particle can knock out many electrons before it loses all of its kinetic energy

also α-particles travel at low speeds, so frequent collision events take place in short range

therefore α-radiation has strong ionising ability but weak penetration power

➢ the table below summarises penetration and ionising power of radioactive radiation

	α-radiation	β-radiation	γ-radiation
penetration power	low stopped by thick cardboard or a few centimetres of air	fair stopped by metal plates of a few millimetres thick	high stopped by thick lead or concrete of a few centimetres
ionising power	high	fair	low

15.2.3 laws of conservation

some of the conserved quantities during any decay process are:

- conservation of *electric charge*
- conservation of *charge number* (Z)

  this is a consequence of the conservation of electric charge
- conservation of *mass-energy*

  Einstein's theory of relativity suggests that mass is equivalent to energy $E = mc^2$

  for any naturally occurring nuclear decay process that releases energy, total mass of product particles will be *less* than that of the original particles

  reduction in mass becomes of the energy released, which can be kinetic energy of product particles, electromagnetic energy of γ-radiation emitted, etc.
- conservation of *mass number* (A)

  notice that mass number is not exactly proportional to mass of nucleus

  mass itself is not conserved during the reaction but only the mass number is conserved
- conservation of *momentum* (review §6.1.3 if needed)

  no external force is involved in nuclear decays, so total momentum is also constant

  more specifically, when a stationary nucleus decays into two product particles, for example, the two product particles after an α-decay process should move off with equal but opposite momenta

15.2.4 decay equations

using nuclide notation, radioactive decay processes can be described by *decay equations*

the conserved quantities provide a simple guideline to write the decay equations

- sum of charge numbers (Z) on both sides of the equation should add up
- sum of mass numbers (A) on both sides of the equation should add up

Example 15.5 The first reported nuclear reaction was credited to Ernest Rutherford. By firing α-particles into pure nitrogen, Rutherford observed the ejection of hydrogen nuclei from the gas, which is now regarded as the discovery of protons. This reaction can be given by:

$$^{14}_{7}\mathrm{N} + ^{4}_{2}\alpha \longrightarrow ^{17}_{8}\mathrm{O} + ^{1}_{1}\mathrm{p}$$ □

Example 15.6 One reaction that is widely used to power nuclear reactors is the induced fission of uranium-235. The reaction can be triggered by bombarding the uranium-235 nuclei with slow neutrons. The uranium nucleus would split up into two lighter nuclei and release a large amount of energy. One of such reactions can be described by:

$$^{235}_{92}\mathrm{U} + ^{1}_{0}\mathrm{n} \longrightarrow ^{141}_{56}\mathrm{Ba} + ^{92}_{36}\mathrm{Kr} + 3^{1}_{0}\mathrm{n}$$ □

equation for α-decays

if a nuclide $^{A}_{Z}X$ undergoes α-decays, then we have:

$$\boxed{^{A}_{Z}X \longrightarrow ^{A-4}_{Z-2}Y + ^{4}_{2}\alpha}$$

the original nucleus lost two protons and two neutrons during the decay

on both sides of the equation, there are Z protons and $(A-Z)$ neutrons

therefore α-decay simply is an α-particle escaping from the nucleus

Example 15.7 The process in which a uranium-238 nucleus naturally decays into a thorium-234 nucleus through α-emission can be written as:

$$^{238}_{92}\mathrm{U} \longrightarrow ^{234}_{90}\mathrm{Th} + ^{4}_{2}\alpha$$ □

equation for γ-decays

γ-radiation has zero charge and zero mass, so decay equation is very straightforward:

$$\boxed{^{A}_{Z}X \longrightarrow ^{A}_{Z}X + ^{0}_{0}\gamma}$$

since γ-radiation is pure energy, so there is no change of structure of the nucleus in any way

problems with β-decays

for β-decays, one might attempt to write: $^{A}_{Z}X \longrightarrow ^{A}_{Z+1}Y + ^{0}_{-1}\beta$

on the left, nuclide X has Z protons and $(A-Z)$ neutrons

but on the right, nuclide Y has $(Z+1)$ protons and $(A-Z-1)$ neutrons

one neutron has transformed into a proton through giving off an electron: $^{1}_{0}\mathrm{n} \longrightarrow ^{1}_{1}\mathrm{p} + ^{0}_{-1}\mathrm{e}$???

but how can neutrons and protons transform into one another?!

further experiments show that β-particles emitted for same source have *a range of speeds*

this seems to indicate that energy released from a β-decay can be indefinite

both considerations imply something is wrong, our understanding of β-decay is not complete

- protons and neutrons cannot be the most fundamental particles of nature

  there must exist some more fundamental constituent particles

  these are now known as *quarks*, we will discuss them in the next section

– there must be some other unseen particles released during β-decays
these particles carry off a fraction of the energy released from the reaction
for conservation laws to hold, they must be chargeless and very light (maybe massless)
these ghostly particles are called **neutrinos**, or more precisely, **anti-neutrinos**

➢ we thereby can rewrite the β-decay equation for nuclide ${}_Z^A X$:

$$\boxed{{}_Z^A X \longrightarrow {}_{Z+1}^{\ \ A} Y + {}_{-1}^{\ 0}\beta + {}_0^0\bar{\nu}} \quad \text{or} \quad \boxed{{}_0^1\mathrm{n} \longrightarrow {}_1^1\mathrm{p} + {}_{-1}^{\ 0}\beta + {}_0^0\bar{\nu}}$$

➢ distribution of kinetic energy of β-particles can be then explained
only total energy of β-particle and anti-neutrino is constant
anti-neutrinos also carry some energy, so β-particles have a range of energies

Example 15.8 Radon ${}_{86}^{222}\mathrm{Rn}$ decays in a sequence of processes to form bismuth ${}_{83}^{214}\mathrm{Bi}$ by emitting α-particles and β-particles. For the decay chain of each radon nucleus, how many α-particles and β-particles are emitted?

✍ let x be number of α-particles emitted and y be number of β-particles emitted

$${}_{86}^{222}\mathrm{Rn} \longrightarrow {}_{83}^{214}\mathrm{Bi} + x\cdot{}_2^4\alpha + y\cdot{}_{-1}^{\ 0}\beta + y\cdot{}_0^0\bar{\nu}$$

mass number and charge number are conserved: $\begin{cases} 4x + 214 = 222 \\ 2x - y + 83 = 86 \end{cases} \Rightarrow \quad x = 2,\ y = 1$ □

15.3 fundamental particles

In early 20th century, lots of new subatomic particles were discovered in *cosmic rays* and *particle accelerators*. Many of these particles did not fit into the model where proton, neutrons and electrons are the fundamental building blocks for the physical world. To incorporate these subatomic particles discovered at the time, physicist worked out the **Standard Model of particle physics**.[104]

[104]The theory of the Standard Model was developed in stages since the 1950s, through the work of many great scientists, with the current formulation being finalized in the 1970s.

– *Chen Ning Yang* and *Robert Mills* developed the concept of *gauge theory*, to provide an explanation for the interaction between elementary particles.
– *Sheldon Glashow* proposed the symmetry group that forms the basis of the accepted *electroweak theory*, in which the electromagnetic interaction and weak interactions are unified into a single force.
– *Peter Higgs* and two other groups proposed the *Higgs mechanism* that give rise to mass generation for elementary particles without violating gauge theory through a process called *symmetry breaking*.
– *Steven Weinberg* and *Abdus Salam* incorporated the Higgs mechanism into Glashow's theory, finalizing

15.3.1 particles & anti-particles

for any particle p, there exists an associated **anti-particle** $\bar{\text{p}}$

➢ an anti-particle has the same mass but opposite charge to its counterpart

➢ when a particle meets its anti-particle, they **annihilate** each other and produce two γ-photons

combined mass-energy of the pair is converted into electromagnetic energy [105]

Example 15.9 Suggest the properties of the anti-particle of the electron.

✍ anti-particle of electron has same electron mass: $m_e = 9.11 \times 10^{-31}$ kg

but it has a positive charge: $q = +e = +1.60 \times 10^{-19}$ C

anti-particle of electron is usually called a **positron** (positive electron) with the symbol e^+ □

15.3.2 quarks & hadrons

one type of the elementary matter particle is the **quark**

➢ quarks come in six varieties, or six **flavours**

they are **up** (u), **down** (d), **strange** (s), top (t), bottom (b), charm (c)

in A-Level exams, you only need to know up (u), down (d) and strange (s) quarks

➢ each quark carries an electric charge, as given in the table below

quark	u	d	s	t	b	c
charge	$+\frac{2}{3}e$	$-\frac{1}{3}e$	$-\frac{1}{3}e$	$+\frac{2}{3}e$	$-\frac{1}{3}e$	$+\frac{2}{3}e$

note that all quarks carry a fraction of the elementary charge unit

➢ quarks can combine to form composite particles called **hadrons**

there are two ways that several quarks can make up a hadron

---

the unified electroweak theory.

– *David Gross*, *Frank Wilczek*, *David Politze*, and many others developed *quantum chromodynamics*, the theory of the strong interaction, into its modern form.

– …

Being the most accurate accurate scientific theory known to human beings, the Standard Model is now regarded as one of the greatest triumphs of modern physics. The theory not only describes the particles known to scientists, but also predicted new particles, including the *Higgs boson*.

[105]The energy release from the annihilation of particle pairs could be potentially used a energy source. The idea of using antimatter to power spaceships or weapons can be found in many science fiction stories.

- hadrons can be made up of three quarks (qqq), called **baryons**

  members of the baryon family include protons and neutrons

  a proton consists of two up quarks and one down quark (uud)

  a neutron consists of one up quark and two down quarks (udd)

- hadrons can also be made up of one quark and one anti-quark (q$\bar{\text{q}}$), called **mesons**

  you don't need to memorise any example of meson

  you are only required to identify if a particle is a meson given the quark composition

➢ quarks and hadrons are affected by the **strong nuclear force**

strong nuclear force is a very short-ranged attractive force

- it is responsible for hold quarks close together to from hadrons
- it is also responsible for the attraction between hadrons

  for example, protons and neutrons are held together in a nucleus by strong force [106]

➢ quarks only exist in hadrons, i.e., there are no single free quarks in nature[107]

Example 15.10 By reference to the quark composition, explain the electric charge of protons and neutrons.

✍ charge of proton (uud): $q_{\text{p}} = 2q_{\text{u}} + q_{\text{d}} = 2 \times \left(+\frac{2}{3}e\right) + \left(-\frac{1}{3}e\right) = +e$

charge of neutron (udd): $q_{\text{n}} = q_{\text{u}} + 2q_{\text{d}} = \left(+\frac{2}{3}e\right) + 2 \times \left(-\frac{1}{3}e\right) = 0$ □

Example 15.11 A meson has an electric charge of $+e$ and is known to contain an up quark. Determine a possible flavour of the other quark.

✍ charge of the other quark is: $(+e) - \left(+\frac{2}{3}e\right) = +\frac{1}{3}e$

this could be an anti-down ($\bar{\text{d}}$) or an anti-strange quark ($\bar{\text{s}}$) □

15.3.3 leptons

another type of the elementary matter particle is the **lepton**

➢ leptons are not affected by strong nuclear forces

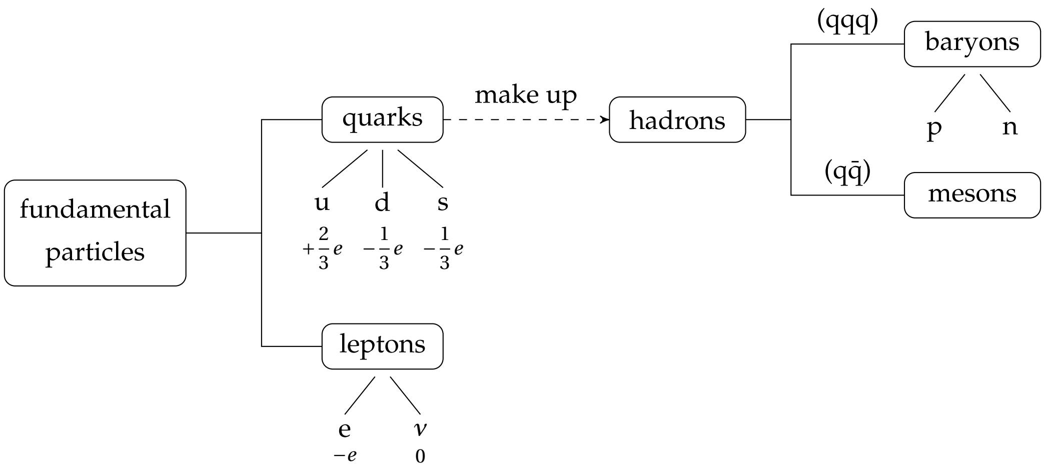
---

[106]For a small nucleus, strength of strong nuclear force is way greater than the electric repulsion between positively-charged protons, hence the force bind protons and neutrons together forming a stable nucleus. However, since the strong nuclear force has a very short range, the repulsive electrostatic force between the protons might dominate the attractive nuclear force for an over-sized nucleus. Such large nuclei become unstable and are likely to undergo *radioactive decays*.

[107]This is known as the *asymptotic freedom* of strong nuclear forces.

- members of the lepton family include electrons (e), neutrinos (ν), muons (μ), taons (τ)
  in A-Level exams, you only need to know electrons (e) and neutrinos (ν)
- for each particle, one can assign a **lepton number** L
  - a lepton has lepton number $L = +1$
  - an anti-lepton has lepton number $L = -1$
  - any other particle has lepton number $L = 0$

classification of sub-atomic particles within the A-Level syllabus[108] is summarised below



a coarse guide for classifying sub-atomic particles in A-Level physics

15.3.4 β-decays

recall positron, the anti-particle of electron, so there are two types of β-particles: β^- and β^+
hence two types of β-decays are possible: β^--decay and β^+-decay

β^--decay revisited

a neutron changes into a proton during a β^--decay: $\boxed{{}^{1}_{0}\text{n} \longrightarrow {}^{1}_{1}\text{p} + {}^{0}_{-1}\beta + {}^{0}_{0}\bar{\nu}}$
we have learned about the quark structures of protons (uud) and neutrons (udd)
therefore β^--decay process at a more fundamental level is: $\boxed{\text{d} \longrightarrow \text{u} + \beta^- + \bar{\nu}}$

[108]This is an over-simplified version of the Standard Model. I only included the particles that you need to know for the A-Level exams. There are a lot of missing pieces that are way beyond the scope of our course.

β^+-decay

β^+-decay occurs when a proton changes into a neutron and emits a positron

β^+-decay process is described by the equation: $\boxed{{}^1_1\mathrm{p} \longrightarrow {}^1_0\mathrm{n} + {}^{\ 0}_{+1}\beta + {}^0_0\nu}$

in terms of quarks, β^+-decay can be rewritten as: $\boxed{\mathrm{u} \longrightarrow \mathrm{d} + \beta^+ + \nu}$

- nature of β-decay processes is the transformation of quark flavours

  the interaction responsible for this transformation is the **weak nuclear force**
- there is yet another law of conservation – the **conservation of lepton number**

  we can use lepton number to predict whether a neutrino or an anti-neutrino is emitted

Example 15.12 Verify the conservation of (a) charge number, (b) mass number, and (c) lepton number for the β^+-decay process: ${}^1_1\mathrm{p} \longrightarrow {}^1_0\mathrm{n} + {}^{\ 0}_{+1}\beta + {}^0_0\nu$

✍ charge number conservation: $+1 = 0 + (+1) + 0$ ✓

mass number conservation: $1 = 1 + 0 + 0$ ✓

lepton number conservation: $0 = 0 + (-1) + (+1)$ ✓

so this reaction does not violate any of these conservation law □

Example 15.13 *Electron capture* is a process where an electron in an atom's inner shell is drawn into the nucleus and combines with a proton. The result is to form a neutron and a neutrino. The process can be given by: ${}^1_1\mathrm{p} + {}^{\ 0}_{-1}\mathrm{e} \longrightarrow {}^1_0\mathrm{n} + +{}^0_0\nu$. Show that the process satisfies the conservation of (a) charge number, (b) mass number, and (c) lepton number.

✍ charge number conservation: $+1 + (-1) = 0 + 0$ ✓

mass number conservation: $1 + 0 = 1 + 0$ ✓

lepton number conservation: $0 + 1 = 0 + 1$ ✓

so this reaction does not violate any of these conservation law □

Index